INTERNATIONAL ECONOMIC REVIEW

January, 1961
Vol. 2, No. 1

CONTENTS

The Problem of the Balance of Payments — *I. F. Pearce* 1

Efficient International Trade and Equalization of Factor Prices — *Stanley Reiter* 29

On Pure and Mixed Statistical Estimation in Economics — *H. Theil and A. S. Goldberger* 65

A Schema for the Supply Theory of Money — *Karl Brunner* 79

Profit, Output, and Liquidity in the Theory of Fixed Investment — *Murray Brown* 110

On the Formation of Prices — *Takashi Negishi* 122

Production Functions with Restricted Input Shares — *P. K. Newman and R. C. Read* 127

Published by
THE KANSAI ECONOMIC FEDERATION, OSAKA

INTERNATIONAL ECONOMIC REVIEW

Published by the Kansai Economic Federation with
the Collaboration of the Wharton School of Finance and Commerce,
University of Pennsylvania, and the Institute of Social and
Economic Research, Osaka University.

CHIEF EDITOR
L. R. KLEIN
Department of Economics,
Wharton School of Finance
and Commerce, University
of Pennsylvania,
Philadelphia 4, Pa.,
U.S.A.

CO-EDITOR
M. MORISHIMA
The Institute of Social
and Economic Research,
Osaka University,
Toyonaka, Osaka,
Japan

ASSOCIATE EDITORS:

M. ALLAIS (École Nationale Supérieure des Mines de Paris, France)
T. M. BROWN (Royal Military College, Canada)
G. DEBREU (Yale University, U.S.A.)
R. M. GOODWIN (Peterhouse, Cambridge, U.K.)
F. H. HAHN (Churchill College, Cambridge, U.K.)
B. HANSEN (Konjunkturinstitutet, Sweden)
L. HURWICZ (University of Minnesota, U.S.A.)
D. PATINKIN (The Hebrew University, Israel)
P. NØRREGAARD RASMUSSEN (University of Copenhagen, Denmark)
N. S. R. SASTRY (Reserve Bank of India, India)
R. H. STROTZ (Northwestern University, U.S.A.)
T. W. SWAN (Australian National University, Australia)
H. THEIL (Netherlands School of Economics, Holland)

BUSINESS OFFICE OF THE INTERNATIONAL ECONOMIC REVIEW: Kansai Economic Federation,
Osaka Building, Soze-cho, Kita-ku, Osaka, Japan.

Copyright © 1961 by Kansai Economic Federation.

Printed by Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.),
No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

INTERNATIONAL ECONOMIC REVIEW

VOLUME 2
1961

Published by
THE KANSAI ECONOMIC FEDERATION, OSAKA

Printed by Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.)
No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

INTERNATIONAL ECONOMIC REVIEW

January, 1961
Vol. 2, No. 1

THE PROBLEM OF THE BALANCE OF PAYMENTS*

BY I. F. PEARCE

A REVIEWER of Professor J. E. Meade's massive work, *The Theory of International Economic Policy*, has commented that since everything is fully worked out therein, we should now be spared the large number of articles we might otherwise have had to read. The present writer is sufficiently in sympathy with all that is implied by this observation to have hesitated before preparing for publication a paper on a subject which Professor Meade has treated. On the other hand, intermittent manipulation over the past few years of a six-good model which is in some ways more sophisticated than that upon which Professor Meade's argument is based has led to the conviction that there is still something worthwhile to be said.

This model, introduced below, is designed to show what precisely is required to correct an adverse trade balance when there exists a class of goods excluded from trade by tariffs and transport costs, given that full employment is to be maintained continuously both at home and abroad. The formal results obtained are not recognizably equivalent to those given by Professor Meade,[1] primarily because Professor Meade makes use of an unusual and somewhat complicated definition of elasticity of supply. This point is further explained below. A more conventional definition yields results which, although still complex, are a great deal simpler and more readily interpretable than Professor Meade's. Greater simplicity makes it possible to introduce more realistic assumptions to the point where almost everything which has been said about the problem can be summarized in three algebraic formulas. In particular, the model allows for some home production of the class of goods imported, *a feature which permits us without absurdity to define two trading countries as one country (small) and*

* Manuscript received May 16, 1960, revised August 24, 1960.

[1] J. E. Meade, *The Balance of Payments: Mathematical Supplement* (London: Oxford University Press, 1951), Table X, p. 156.

the rest of the world. The so-called "elasticity" and "absorption"[2] approaches are fully integrated.

One must also recognize that, despite a great deal of discussion, there still remains considerable apparent disagreement on some points. Professor Machlup, for example, has recently listed[3] no less than fifteen papers and books written between 1945 and 1955, each of which seems to reverse the conclusions set out in its immediate predecessor. The model below is sufficiently general to include as special cases many of those referred to by Professor Machlup. Seemingly conflicting arguments can thus be reconciled.

Again, results obtained suggest that although most effort seems so far to have been directed towards discovering what change in the real terms of trade must be achieved to correct a trade deficit, this is probably of less importance than the need to reduce the price of non-traded goods relative to traded goods in general. Explicit formulas for all necessary relative price changes reveal this clearly. *Indeed, it may well be that currency depreciation succeeds in practice more because it decreases the price of non-traded goods relatively than because it affects the real terms of trade.*

Finally, it is emphasized that a rise in import prices and a cut in export prices can never by itself reduce real consumption sufficiently to correct a trade deficit, as long as full employment is to be maintained. There must be a cut in spending approximately equal in real terms to the original trade deficit *in addition to* the real terms of trade effect.

Part 1 of this paper is introductory. Its purpose is solely to clear up certain ambiguities of terminology and objectives. In Part 2 a simplified model is developed for the purpose of exposition only. The fully general results are obtained in Part 3. This section may be omitted by readers not interested in proofs. In Part 4 results are restated so as to show the successive effects of the introduction of more realistic assumptions. Qualitative conclusions are drawn and implications discussed.

1. TERMINOLOGY AND OBJECTIVES

Relative or absolute approach? In setting up a model to illustrate the economic effects of any kind of policy change there is always one

[2] S. S. Alexander, "Effects of a Devaluation on a Trade Balance," *International Monetary Fund Staff Papers*, Vol. II (1951-52), pp. 263-78.

[3] F. Machlup, "The Terms-of-trade Effects of Devaluation upon Real Income and the Balance of Trade," *Kyklos*, Vol. IX (1956).

important choice to be made. Either some one good must be taken as a *numéraire*, in which case all prices are relative and expenditure is measured in units of the *numéraire*, or the absolute change in at least one variable must be specified beforehand. This follows because, on the most usual assumptions, doubling all prices and expenditures has no effect upon the real situation. There is always a unique real equilibrium, but an infinity of possible equilibria in terms of absolute prices.

In the international trade context this problem becomes especially important, for it means that the effect of a currency depreciation can always be offset, or more than offset, by a general inflation or deflation in one or other country. Moreover, any policy designed to correct a balance of payments deficit is precisely the policy which calls for some change in overall consumption. And where a situation demands *both* a general consumption cut *and* relative price changes everywhere, the very notion of inflation and deflation becomes ambiguous. The precise meaning of a statement like "the effect of currency depreciation on the balance of payments" is far from obvious. Such an "effect" will be different according to the measure of inflation which is assumed to accompany the depreciation and will be different again according to the definition of "inflation" chosen.

In most models it has been usual to assume that *in fact* one price will remain fixed—usually the price of exportables. Where there exists a class of goods excluded from trade, this is clearly unsatisfactory. Alternatively, Professor Meade has chosen the apparently more attractive assumption that it is the money wage rate which does *in fact* remain constant when currency is depreciated. In this connection it is important that we should understand two things.

First, for reasons given above, it is clear that if we are looking at an algebraic formula measuring the effect of some policy expressed in absolute terms (say currency depreciation), then the form of that formula depends precisely on the arbitrary choice of price to be kept constant. There are an infinity of possible formulas which energetic students of economics might devise. For any one of these to be of use to the policy maker it must be assumed either:

(a) That monetary and fiscal policy can be applied with a very remarkable degree of finesse, and that the price to be held constant can be unambiguously defined and identified in practice or,

(b) That the formula presented is derived from a very complex dynamic model with a high degree of plausibility, so that the

precise order of, and time lags in, successive price adjustments are known for a certainty.

The second point to be made is that in a general static model there is no real case for any one price to be preferred as an anchor rather than another. The attractive fixed money wages assumption demands, unfortunately, the further assumption that only one homogeneous factor of production exists, namely labor. Even if this were realistic, it would still appear doubtful to assume that wage rates will remain constant where, as part of the working of the model, labor is supposed to be attracted from one use to another by a price mechanism.

It need hardly be stressed also that in the real world we do observe workers in export industries gaining or losing in relation to others. It is easy to show on theoretical grounds that the magnitude of these changes cannot be assumed to be insignificant in relation to the relative commodity price changes in which we are interested. There is more than one kind of labor and more than one factor of production.

On the other hand, any one of the infinity of possible formulas expressed in absolute terms is easily converted into a unique statement in relative price terms. For example, Professor Meade's formula (*t*), *ioc. cit.*, for the effect of depreciation on the trade balance divided by the corresponding formula (*v*) for its effect on the real terms of trade at once yields a further formula which can be interpreted as the effect on the trade balance of a change in the real terms of trade. Our plea is simply that, given our present state of knowledge, it is only in this relative form that any formula can be used. Since this is so, any *numéraire* will do. To attempt to find a result in absolute terms is asking too much of our techniques.

Accordingly, we shall, in what follows, confine ourselves to statements of the following kind: "If the balance of payments is to be improved we require (a) a certain adjustment in the real terms of trade, (b) a certain adjustment in the ratio of prices of traded and non-traded goods in both countries, and (c) an adjustment of so and so much in overall consumption in each country." The symbols for prices and income represent money prices and money incomes, but they appear finally only in ratio form.

The question of how exactly the required adjustments are to be attained is a difficult problem in dynamic economics which is hardly touched upon. Fortunately, where there is an infinity of possible solutions all giving the same real solution, there is equally an infinity of possible policy combinations which might achieve the desired result. Once the static solution is known, trial and error methods, applied in

the light of that knowledge, stand a good chance of success. We comment very briefly upon the part which currency depreciation might play in the process.

The absorption problem. It is a fundamental assumption of what follows that total demand is always equal to total supply both before and after any policy change. By "total" we mean foreign and home together. All that we imply by this is that in any period in which we observe a trade deficit the price of every product is such that the world market is just, but only just, cleared. The consequence is that the identity

$$\text{Consumption} + \text{Investment} = \text{Production} + \text{Imports} - \text{Exports}$$

holds at all times in all countries. If imports do not equal exports, so that a trade deficit exists, this must be financed by a movement abroad of exchange reserves. Because total demand everywhere equals total supply, the situation is one of equilibrium. No economic mechanism will operate to change it as long as no action is taken to check the outflow of reserves. To discover what action is needed to check such an outflow is the object of our enquiry.

Special attention is drawn to this rather ordinary assumption solely because it will be recognized as the starting point of those who have favored the so-called "absorption" approach to the problem of the balance of payments. The fact that the income expenditure identity plays a key role in the work below confirms our earlier claims that absorption and elasticity approaches have been fully integrated.

Problems of aggregation. It will later be necessary to make extensive use of the theory of demand. In particular, the elasticity of demand (ε_{ij}) for the ith good with respect to the jth price will be broken down into substitution elasticity (σ_{ij}) and income effect $-(p_jX_j/p_iX_i)C_i$.[4] X_i is the quantity consumed of the ith good and p_i is its price. C_i is the marginal propensity to consume the ith good. We shall presume also that the following well known properties of substitution elasticities hold:

(i) $\sigma_{11} + \sigma_{12} + \sigma_{13} = 0$ (for three goods);
(ii) σ_{ii} is negative;
(iii) $\sigma_{ii}\sigma_{jj} - \sigma_{ij}\sigma_{ji}$ is positive;
(iv) $\sigma_{ji} = (p_iX_i/p_jX_j)\sigma_{ij}$.[5]

[4] P. A. Samuelson, *Foundations of Economic Analysis* (Cambridge, Massachusetts: Harvard University Press, 1947), chap. v. Note that σ_{ij} above is p_jK_{ij}/X_i in Professor Samuelson's notation.

[5] Cf. Samuelson, *loc. cit.*

The assumption that community demand elasticities have the same properties as individual demand elasticities requires justification. This problem has been dealt with elsewhere in detail by the present writer.[6] Three points, however, must be made here.

First, any money income redistribution due to relative commodity price changes is assumed to be random with respect to tastes. Roughly speaking, this means that there is no bias towards the rich gaining at the expense of the poor or vice versa. If bias is present, the effect of this must be separately considered.

Secondly, demand for each particular good is written as a function of prices and total *consumption* rather than total income. In these circumstances the breakdown of elasticity into income and substitution effect involves the assumption that changes in the individual's stock of wealth will not affect the shape of his preference pattern or, more strictly, will not affect his marginal rates of substitution between goods currently consumed.[7] We note also that aggregate consumption does not depend on current prices, although individual consumption does. This is apparent because aggregate consumption includes government consumption, which is automatically adjusted so as to maintain full employment at all times.

Finally, that part of investment which represents a demand for goods is included in aggregate consumption. Marginal producers whose investment activity is affected by current prices are regarded as a class of consumers with "indifference curves" defined by production functions. Redistribution of investment expenditure between producers is exactly analogous to the redistribution of consumption expenditure already referred to. Thus, all that we know about the theory of consumer demand applies equally to investment demand, and we remain free to make use of the properties of demand elasticities quoted above.

The fact that goods are classified into exports, imports and non-traded goods rather than into groups with similar characteristics does not create any special difficulty. As long as significant tariff barriers exist, small relative price changes cannot cause any good which was formerly an import to become an export. The fact that relative price changes might cause some good, formerly an import, to be produced entirely at home does not matter as long as the model allows for some home production of the whole class of imports and as long as

[6] Cf. a forthcoming paper entitled "The Theory of Aggregate Demand."

[7] Cf. I. F. Pearce, "Demand Analysis and the Savings Function," *Australian Economic Record*, Vol. XXXIV, No. 67 (April, 1958).

changes are continuous rather than discrete.

Notation. For the benefit of readers who wish to turn at once to the results given in Part 4, an explanation of the notation used is given in this section. Symbols for elasticity (ε), substitution elasticity (σ), and marginal propensity to consume the ith good C_i have already been introduced.

Good 1 is the class of exportables of the home country, good 2 is the class of importables, while good 3 is excluded from trade. X_1, X_2, and X_3 represent the total quantities consumed at home of the three goods. X'_1, X'_2, and X'_3 are the quantities consumed abroad. Primes are used quite generally to denote "belonging to the foreign country" wherever there is ambiguity. Thus C'_3 is the foreign country's marginal propensity to consume non-traded goods. Where there is no ambiguity primes are dropped even though symbols may refer to the foreign country. In some cases primes have been used outside brackets to indicate that all symbols inside the brackets refer to the foreign country.

x_1 is the quantity of good 1 exported from the home country. x_2 is the quantity of good 2 imported by the home country. $(X_1 + x_1)$ therefore is total production of good 1 in the home country and $(X'_1 - x_1)$ is total production of the same good abroad.

As all prices appear ultimately as price relatives, we need only work in units of home currency. p_1, p_2, and p_3 are the absolute home prices of each good. Prices abroad, both absolute and relative, will be different from those at home because of tariffs and transport costs and because some goods are not traded. For foreign prices we use therefore the symbols q_1, q_2, q_3.

Supply is assumed to be a function of all commodity prices. Supply elasticity is written S_{ij}. This measures the percentage change in the output of good i due to a one per cent change in the price of good j. It is assumed that supply functions are homogeneous of order zero in prices; that is, the doubling of all prices would leave supply unchanged. The transformation surface defining optimum production possibilities is supposed also to be concave to the origin.[8] Thus supply elasticities have the following properties analogous to demand elasticities:

(i) S_{ij} is all substitution elasticity; there is no income effect;
(ii) $S_{11} + S_{12} + S_{13} = 0$ (for three goods);
(iii) S_{ii} is positive—the higher the price the greater the supply;

[8] Cf. Samuelson, *loc. cit*, pp. 234–35.

(iv) $S_{ii}S_{jj} - S_{ij}S_{ji}$ is positive;
(v) $S_{ji} = (p_iX_i/P_jX_j)S_{ij}$.

In the following sections these definitions will be repeated where appropriate when the symbol is first introduced to avoid referring back.

2. A SIMPLIFIED EXPOSITORY MODEL

We begin with the simple case where all goods enter into trade. The balance of payments in terms of home currency is

$$B = p_1x_1 - q_2x_2 , \tag{1}$$

that is, receipts from exports less payments for imports. The home country receives only the home price for its exports and pays only the foreign price for its imports. The price ratio q_2/p_1 gives a measure of the real terms of trade.

q_1/p_1 is greater than one by an amount equal to the tariff and transport cost loading when exports are delivered abroad.[9] It is assumed throughout that the loadings q_1/p_1 and p_2/q_2 are constant. Thus, any percentage change in the real terms of trade q_2/p_1 must always be exactly the same as the percentage change in either the ratio p_2/p_1 or the ratio q_2/q_1. We shall, in fact, use the percentage change in p_2/p_1 to measure the change in the real terms of trade. The assumption of constant loadings is equivalent to assuming that tariffs are *ad valorem* and that transport costs are negligible or have special properties. It will be seen later that to make loadings variable would introduce additional complications without demonstrating any new principle.

The balance of payments (equation (1)) is now written in units of good 1 instead of money:

$$B/p_1 = x_1 - (q_2/p_1)x_2 . \tag{2}$$

Differentiating (2) totally we derive an expression for β, the change in the balance due to a one per cent change in the real terms of trade,

[9] The existence of a commodity "transport" raises an insuperable aggregation problem in a six-good model such as the present one. The reader may prefer, therefore, to assume away transport costs and attribute all loading to tariffs. It would be tedious, but not difficult, to distinguish between home produced importables and importables with a fixed proportion of transport from abroad so as to avoid contradiction. But this is not attempted here, as it is considered that the disadvantage of greater complication would not be outweighed by corresponding advantages.

$$(3) \qquad \beta = \frac{p_2}{p_1}\frac{dx_1}{d(p_2/p_1)} - \frac{q_2}{p_2}\left(\frac{p_2}{p_1}\right)^2 \frac{dx_2}{d(p_2/p_1)} - \frac{p_2}{p_1}\frac{q_2}{p_2}x_2 .$$

This is obtained by writing q_2/p_1 as q_2p_2/p_2p_1 and remembering that q_2/p_2 is constant by virtue of the constant tariff loading assumption.

To avoid unnecessary complication we now suppose that trade is initially balanced. It would not be difficult to work the general case, but we lose very little by the restriction, since in most practical cases any imbalance in trade will not be allowed to grow large relative to amounts traded. Thus

$$p_1x_1 = q_2x_2 ,$$

so that β reduces to

$$(4) \qquad \beta = x_1\left(\frac{p_2}{p_1}\frac{dx_1}{d(p_2/p_1)}\frac{1}{x_1} - \frac{p_2}{p_1}\frac{dx_2}{d(p_2/p_1)}\frac{1}{x_2} - 1\right) .$$

Now the form of the expression $\left(\frac{p_2}{p_1}\frac{dx_1}{d(p_2/p_1)}\frac{1}{x_1}\right)$ tempts us to call it the elasticity of demand, for it looks like the percentage change in x_1 due to a one per cent change in the "price" p_2/p_1. Moreover, a one per cent change in p_2/p_1 can always be interpreted as a one per cent fall in p_2 with p_1 constant or a one per cent rise in p_1 with p_2 constant. Thus, it might be argued that (4) can be written

$$(5) \qquad \beta = -x_1(E_1 + E_2 + 1) ,$$

where E_1 and E_2 are elasticities of demand. It is in this form that the effect of devaluation on the balance of payments is often expressed.

Unfortunately this "solution" to the problem does not tell us very much. This is because the expressions we have written as E are not really elasticities in the ordinary sense, but the total effect on x both of price changes and any other subsidiary changes which may be appropriate. We can infer nothing about either their sign or their magnitude. Moreover, it is doubtful if they have yet been satisfactorily defined in terms of more familiar parameters.

Elasticity of demand ought only to mean the percentage change due to a one per cent change in price, all other prices and *total spending* constant. Only on this definition can elasticity be thought of as a parameter the value of which can, in principle, be known independently of the problem under review. Even in the simplest model E cannot mean this; if the balance of payments is really changed while production remains constant, spending *must* have changed. The dx_1 of (4) will include not only the change due to the change in price p_1

but also some further change due to the change in spending. Moreover, x_1 is not necessarily the total demand for good 1 abroad. Some part of this may be met out of production abroad. Any price induced change in this supply will affect the demand for x_1 for yet another reason.

E is, in fact, what has been called elsewhere a "total" elasticity.[10] Such an elasticity is meaningless until its elements are specified. Indeed, since in the present context supply and demand are always equal, E could just as easily be called supply elasticity. The reader will note that in deducing formula (5) we took no account of supply conditions, nor did we at any time make use of the assumption that only two kinds of goods exist. (5) is, in fact, an alternative way of writing any possible formula which might be devised to cover any kind of case. By appropriately defining E, (5) can be made equivalent to any of the various formulas given below or referred to. It even covers the case where employment is allowed to vary.[11] To obtain a meaningful result it is necessary to define E carefully in terms of the parameters of demand and supply equations and substitute the result into (5). This is what we now proceed to do.

We distinguish first between the demand for exports (x_1) and the total demand for good 1 (exportables) at home (X_1) and abroad (X'_1). The theory of demand tells us that with only two goods total foreign demand for good 1 may be written as a function (f) of the relative price q_2/q_1 and total consumption measured in units of good 1, namely C/q_1. Foreign demand for exports from the home country is total foreign demand less own supply. Thus

$$x_1 = f(q_2/q_1, C/q_1) - g(q_2/q_1) . \tag{6}$$

Reasons have already been given for putting demand as a function of total consumption rather than income. Supply (g) depends only on relative prices and may be thought of in terms of the familiar transformation curve defining optimum production possibilities. Doubling all prices leaves supply unchanged.

We now ask which of the independent variables change and by how much. To find the change in C/q_1 we recall the basic absorption condition

[10] Cf. I. F. Pearce, "Total Demand Curves and General Equilibrium," *Review of Economic Studies*, Vol. XX (3) (1952–53), pp. 216–27.

[11] Cf. John Spraos, "Stability in a Closed Economy and in the Foreign Exchange Market and the Redistribution Effect," *Review of Economic Studies*, Vol. XXIV (1956–57), footnotes 1, 2, and 3.

Consumption + Investment = Production + Imports − Exports ,

quantities being valued at internal prices.
Remembering that we are now looking at the foreign country we have

$$C/q_1 \equiv O/q_1 + x_1 - (q_2/q_1)x_2 \qquad \text{(where } O = \text{production)}$$
$$\equiv O/q_1 + B/q_1 + (1 - p_1/q_1)x_1 .$$

We recognize the last term as foreign tariff receipts in units of good 1. Thus we obtain

$$d(C/q_1) = d(O/q_1) + d(B/q_1) + (1 - p_1/q_1)dx_1 ,$$

where "d" is read "the small change in." p_1/q_1 is, of course, constant by assumption. But with two goods only

$$O/q_1 \equiv (X_1' - x_1) + q_2/q_1(X_2' + x_2) ,$$

so that

$$d(O/q_1) = d(X_1' - x_1) + q_2/q_1 d(X_2' + x_2) + d(q_2/q_1)(X_2' + x_2) .$$

But if full employment is maintained, the ratio of changes in production $d(X_1' - x_1)/d(X_2' + x_2)$ must be equal to minus the ratio of marginal costs, which under competition will be equal to $-q_2/q_1$. Hence,

$$d(X_1' - x_1) + (q_2/q_1)d(X_2' + x_2) = 0 ,$$

so that

$$d(O/q_1) = d(q_2/q_1)(X_2' + x_2) .$$

Thus,

$$(7) \qquad d(C/q_1) = d(q_2/q_1)(X_2' + x_2) + d(B/q_1) + (1 - p_1/q_1)dx_1 .$$

From (6) we now write an expression for the total elasticity of demand for exports (E_1),

$$(8) \qquad \frac{dx_1}{d(q_2/q_1)} \frac{q_2}{q_1} \frac{1}{x_1}$$
$$= -E_1 = q_2/(q_1x_1)(f_{q_2/q_1} + f_{C/q_1}[d(C/q_1)/d(q_2/q_1)] - g_{q_2/q_1}) ,$$

where f_{q_2/q_1}, f_{C/q_1} and g_{q_2/q_1} are the partial derivatives of the functions f and g. This can be converted into the more familiar elasticity notation as follows.

Thinking of f as a function of q_1, q_2 and C separately rather than as a function of the ratios q_2/q_1 and C/q_1, we note that $q_2 f_{q_2}$ is exactly equivalent to $(q_2/q_1)f_{q_2/q_1}$. This must be so, for even if part of any change in q_2/q_1 were due to some change in q_1, this must be matched

by an equal percentage change in C so as to keep C/q_1 constant. And we know that any given percentage change in both q_1 and C is precisely equivalent to the same (negative) change in q_2 with q_1 and C constant. Thus, in our elasticity notation,

$$(q_2/q_1X_1)f_{q_2/q_1} = \varepsilon_{12} ,$$

where ε_{12} is the ordinary partial cross elasticity of demand. The supply function is open to the same interpretation. We notice also that

$$f_C = f_{C/q_1}\frac{d(C/q_1)}{dC} \equiv f_{C/q_1}/q_1 ,$$

so that f_{C/q_1} is the familiar marginal propensity to consume good 1 (C_1).

Thus (8) can be written using (7) as

$$-E_1 = \frac{X_1}{x_1}\varepsilon_{12} - \frac{X_1 - x_1}{x_1}S_{12} + \frac{C_1}{x_1}\frac{q_2}{q_1}(X_2 + x_2) + \frac{C_1}{x_1}\frac{q_2}{q_1}\frac{d(B/q_1)}{d(q_2/q_1)} - (1 - p_1/q_1)C_1E_1 .$$

If we define

$$\alpha = 1/(1 - C_1(1 - p_1/q_1)) ,$$

this becomes

$$E_1 = -\alpha\left(\frac{X_1}{x_1}\varepsilon_{12} - \frac{X_1 - x_1}{x_1}S_{12} + \frac{C_1}{x_1}\frac{q_2}{q_1}(X_2 + x_2) + \frac{C_1}{x_1}\frac{q_2}{q_1}\frac{d(B/q_1)}{d(q_2/q_1)}\right).$$

It should be noted that α is a multiplier, not very different from unity, which exists only because of the tariff loading. Splitting ε_{12} into substitution elasticity σ_{12} and income effect $-(q_2X_2/q_1X_1)C_1$ and using $q_2x_2 = p_1x_1$, we have

$$E_1 = -\alpha\left(\frac{X_1}{x_1}\sigma_{12} - \frac{X_1 - x_1}{x_1}S_{12} + \frac{C_1}{x_1}\frac{q_2}{q_1}\frac{d(B/q_1)}{d(q_2/q_1)} + \frac{p_1}{q_1}C_1\right).$$

But $\sigma_{12} = -\sigma_{11}$ and $S_{12} = -S_{11}$ with two goods only. We know also that $B/q_1 = (p_1/q_1)B/p_1$ or $d(B/q_1) = (p_1/q_1)d(B/p_1)$, so that

$$(9) \qquad E_1 = \alpha\left(\frac{X_1}{x_1}\sigma_{11} - \frac{X_1 - x_1}{x_1}S_{11} - \frac{p_1}{q_1}C_1 - \frac{C_1}{x_1}\frac{p_1}{q_1}\frac{p_2}{p_1}\frac{d(B/p_1)}{d(p_2/p_1)}\right).$$

By an analogous method we may obtain also

$$(10) \qquad E_2 = \alpha\left(\frac{X_2}{x_2}\sigma_{22} - \frac{X_2 - x_2}{x_2}S_{22} - \frac{q_2}{p_2}C_2 - \frac{C_2}{x_2}\frac{d(B/p_1)}{d(p_2/p_1)}\right).$$

Substitution in (5) yields the desired meaningful formula for the effect

on the trade balance of a one per cent change in the real terms of trade

$$(11)\quad \beta = -x_1\left[\frac{\alpha'\left(\frac{X_1}{x_1}\sigma_{11} - \frac{X_1 - x_1}{x_1}S_{11}\right)' + \alpha\left(\frac{X_2}{x_2}\sigma_{22} - \frac{X_2 - x_2}{x_2}S_{22}\right)}{1 - \alpha' p_1 C_1'/q_1 - \alpha q_2 C_2/p_2} + 1\right].$$

At this point it is convenient to make one or two observations of a general nature. First, the formulas (9) and (10), even with the present simplified two-good model, already give some hint of the complexities which can be covered up by the use of vaguely defined expressions such as "the elasticity of demand for imports," for this is precisely what the total elasticities E_1 and E_2 are. On the other hand, it is clear from our results that much of what has been said about the factors governing the magnitude of the "elasticity of demand for imports" by writers who favor the literary intuitive approach is substantially justified. The fact that E is made up of the sum of the Marshallian elasticity and the elasticity of supply simply gives expression to the intuitively obvious idea that a rise in the price of importables will reduce the quantity bought abroad still more when the same price rise serves also to increase the quantity produced at home. Again, the fact that the substitution part of the Marshallian elasticity has a multiplier X/x shows how E will tend to be larger if the importing "country" produces a high proportion of its total consumption of importables. If, for example, the importing country is the "rest of the world," X may be very large relative to x. In this way our formula confirms the idea we have in mind when we say that the price of the exportables of some small region is determined by the world market, and hence elasticity of demand for them is large. Finally, the term in E involving the change in the balance of payments is the much discussed absorption effect. All of these features, with added complications not intuitively obvious, reappear in the three-good model of the next two sections.

Turning now to the formula for β we note:

(i) Our result is totally unlike the celebrated "four elasticities" formula devised by Mrs. Robinson.[12] There are two possible reasons for this, depending upon the way in which Mrs. Robinson's model is interpreted. Either the "four" elasticities are total elasticities not precisely defined, or it must be implicity assumed that demand is independent of income. In the first case we have the difficulty already

[12] Joan Robinson, "The Foreign Exchanges," *Essays in the Theory of Employment* (Oxford: Blackwell, 1947), pp. 134–55.

noted in our comments on (5) above. The second interpretation is no more satisfactory.[13]

(ii) If we assume that there is no home production in either country of the class of goods imported, so that $X_1' = x_1$, $X_2 = x_2$, and $S_{11} = S_{22} = 0$; and if we assume also that there are no tariffs or transport costs, so that $\alpha' = \alpha = 1$, $p_1 = q_1$, and $p_2 = q_2$, then our formula for β becomes

$$\beta = -x_1[(\sigma_{11} + \sigma_{22})/(1 - C_1' - C_2) + 1] \tag{12}$$

which, when we recall that $\sigma_{11} = -\sigma_{12}$ and $\sigma_{22} = -\sigma_{21}$, is the same as that given by Professor Meade for this special case.[14] By recombining income and substitution effects into elasticities ε, β may be rewritten as

$$\beta = -x_1(\varepsilon_{11} + \varepsilon_{22} + 1)/(1 - C_1' - C_2) .$$

Professor Meade accordingly reconciles his formula with the "familiar rule" that devaluation will always improve the balance of payments, if the sum of elasticities is numerically greater than unity, by assuming that $1 - C_1' - C_2$ is always positive. This assumption is crucial and we shall need to discuss it at length. Before turning to this, however, we note in passing one of the curious results of the use of ambiguously defined total elasticities. Professer Meade's fundamental assumption is that supply elasticities are zero. On this assumption he arrives at the "familiar rule." On the other hand, to reduce Mrs. Robinson's formula to the "familiar rule" it is necessary to assume that supply elasticities are infinite. The reader will at once see the reason for this anomaly if the invitation of footnote 13 is accepted.

Consider now the assumption that $1 - C_1' - C_2$ is positive. This is justified by Professor Meade on the ground that it is unlikely in practice that the marginal propensity to import of any country will

[13] Professor Meade in reworking the Robinson formula favors interpretation two (cf. *The Balance of Payments: Mathematical Supplement* chap. x). Readers are invited to attempt a restatement of Meade's mathematics, writing imports as a function of relative prices rather than absolute prices in equations (4.1) and (4.2) (*loc. cit.*) From (5.1) and (5.4) we then discover that equilibrium is impossible unless elasticities of supply and demand are equal but of opposite sign. Making use of this fact it is easy to reduce the four elasticities formula to (5) of the text above.

If the "total" elasticity interpretation is accepted, it is hard to see how elasticity of supply can ever be different from elasticity of demand (except for the fact that different currencies are involved). Once again the four elasticities formula reduces to (5). That it is correct is not disputed. We question only its usefulness and warn that the elasticities which appear in it must not be interpreted as parameters independent of the problem.

[14] *The Balance of Payments: Mathematical Supplement*, Table IV (b), p. 150. Cf. also p. 46, (1).

be greater than one-half. This seems a plausible enough assumption when we look at the real world. Unfortunately, the model to which Professor Meade applies it is a two-good model which, in this respect, is not descriptive of the real world. With two goods only, $C_1 + C_2 = 1$, so that

$$1 - C_1' - C_2 = C_2' - C_2 = C_1 - C_1' .$$

To assume that $1 - C_1' - C_2$ is positive is therefore exactly equivalent to assuming that the home country's marginal propensity to consume its exportables (C_1) is always greater than the foreign country's marginal propensity to consume the same class of goods (C_1'). In this guise it looks much less plausible. It would be surprising, for example, if West Africa's marginal propensity to consume ground nuts and palm oil were greater than that of the rest of the world. Moreover, the sign and magnitude of $C_1 - C_1'$ is the crucial determinant of both the sign and magnitude of β. This may most conveniently be seen by looking at Professor Meade's formula in the form of (12). $\sigma_{11} + \sigma_{22}$ is bound to be negative, and the sum of two substitution elasticities is almost certain to be numerically greater than the difference between two marginal propensities to consume the same good. Indeed, it will be argued in Part 4 that the smaller the substitution effect the higher the degree of complementarity, i.e., the greater the tendency to consume exportables and importables in fixed proportions. Furthermore, if goods must be consumed in fixed proportions, it seems more likely that these fixed proportions will be the same in any two countries, so that C_1' and C_1 tend to be equal. Hence, the smaller is $C_1 - C_1'$. It follows that the sign of β may be positive or negative and, if $C_1 - C_1'$ is very small, β may be very large. In its new form Professor Meade's model suggests that there may be just as many cases where a favorable change in the terms of trade is required to correct a trade deficit as there are cases where the necessary change is adverse. Moreover, such a result is hardly surprising; it is very well known and has been much debated in connection with the transfer problem, which is, in a sense, the same problem in reverse.[15]

Exactly the same considerations apply in the more general case, equation (11). Since α, α', p_1/q_1, and q_2/p_2 are all positive and not very different from unity, the denominator of the first term of (11) has the same characteristics as before. The numerator is again bound

[15] Cf. P. A. Samuelson, "The Transfer Problem and Transport Costs," Parts I and II, *Economic Journal*, Vol. LXII (June, 1952), pp. 278-304, and Vol. LXIV (June, 1954), pp. 264-89.

to be negative but numerically larger by the addition of supply elasticities and by multiplication by X/x etc., as already indicated.

It may be objected, of course, that if Professor Meade's qualitative conclusions are unacceptable because they are drawn from a two-good model in a three-good world, then any conclusions we draw from the same model in any other form must be equally unacceptable. Such objections would be valid. *Everything depends upon the way in which the model generalizes.* If the three-good model contains $1 - C_1 - C_1'$ in the denominator, then Professor Meade's empirical judgment is justified. If, on the other hand, the denominator appears as $C_1 - C_1'$, then the new view is more appropriate. It will be shown in Part 3 that the new view is nearest the mark.

At this point we introduce briefly a diagram which, although not new, serves to bring out certain dynamic aspects of the problem which cannot be treated algebraically and which lead up to a point important to the general case.

Figure 1 represents the familiar Edgeworth box. OB and PB measure the (fixed) quantities $(X_1 + x_1)$ and $(X_2' + x_2)$ of commodities

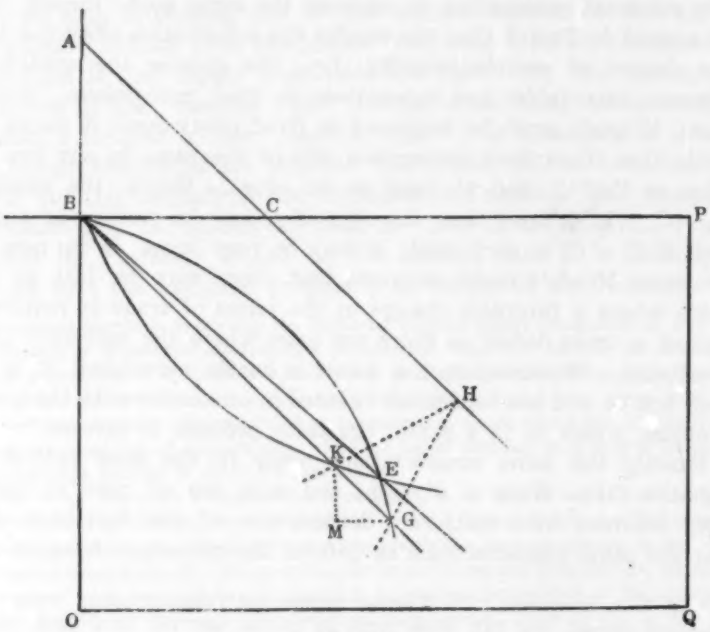

FIGURE 1

produced. *BKE* and *BEG* are "offer" curves cutting at *E*. As drawn, demand elasticities are clearly high.

With terms of trade initially at *BE* and equilibrium consumption *E*, trade is balanced. If country 1 depreciates its currency it will, on impact, worsen its real terms of trade, say to *BG*. The new patterns of demand are defined by the points *K* and *G*. This does not mean that exports now exceed imports. It means only that the *demand* for exports exceeds the *demand* for imports. We have no means of knowing whether these demands will be met or not; it is clear that an overall surplus of demand of *KM* has developed for the product of country 1 and an overall deficit of *MG* for the product of country 2.

One thing, however, we do know. If the object of country 1's depreciation is to build up an export surplus, there is no point in allowing the effect of depreciation to be nullified by a rise in prices due to the inflationary pressure set up thereby. The next logical step is to reduce total spending. Acting alone country 1 could always reduce home demand so as to equate the total demand for its product with supply, that is, until home demand is reduced by an amount *KM*. But such action will also reduce the demand for country 2's product at a time when it is already experiencing deflationary pressure. If, as a result, spending in country 2 is increased, this will recreate the inflationary pressure in country 1. As the diagram is drawn, this process will continue indefinitely. No equilibrium is possible. The increase in the demand for country 1's product due to increased spending abroad is always greater than the decrease due to its own deflation because $C_1 < C_1'$. The object of devaluation is defeated by attempts to maintain full employment.

In fact, the only possible position where total supply is everywhere equal to total demand, with terms of trade *BG*, is at *H*. If such a position were ever reached, depreciation would have created a deficit in the trade balance instead of the surplus intended. Total spending in country 1 is *OA*. The trade deficit is *BA*. More is being consumed than is being produced, and the excess is paid for by a loss of reserves. Country 2 is consuming less than it produces but is gaining reserves. The consumption patterns defined by the point *H* just exhaust total production.

The implication of all this is that for country 1 to create a surplus it should have *appreciated* to improve its terms of trade and *deflated* to cut consumption. This holds as long as marginal propensities to consume take values as drawn, namely, as long as $C_1 < C_1'$. We have no reason to suspect that this case is at all unusual.

To put it the other way round, if two countries found themselves

in situation H with total supply equal to total demand but with country 1 losing reserves, it would be necessary for country 1 to appreciate until the real terms of trade were improved to BE and then to cut spending in order to correct the deficit. This is the case even though demand elasticities are high.

We must emphasize, however, that all this is true only for the simple model. It will become clear later that in a six-good world the dynamics are much more complex. In general it will still be necessary to improve (or worsen) the terms of trade and to cut spending. But the most important price change is usually a rise in the price of traded relative to non-traded goods. The reason why in practice devaluation seems to have been universally successful in correcting balance of payments disorders may be that it achieves the third of these objectives rather than either of the first two.

3. GENERALIZATION OF RESULTS

In this section we introduce the third good and derive the most general expression for β. A parallel formula is obtained also to show what change in the price ratio p_3/p_1 is required to achieve equilibrium.

Proceeding exactly as in Part 2 we break down the total elasticity of formula (5) into its constituent parts, remembering to include the effect of the change in the additional price relative p_3/p_1. Thus we have

$$-E_1 = \frac{X_1}{x_1}\varepsilon_{13} - \frac{X_1 - x_1}{x_1}S_{13} + \left(\frac{X_1}{x_1}\varepsilon_{13} - \frac{X_1 - x_1}{x_1}S_{13}\right)\frac{q_3}{q_2}\frac{d(q_3/q_1)}{d(q_2/q_1)}$$

$$+ \frac{C_1}{x_1}\frac{q_2}{q_1}(X_2 + x_2) + \frac{C_1}{x_1}\frac{q_3}{q_1}\frac{d(q_3/q_1)}{d(q_2/q_1)}X_3 + \frac{C_1}{x_1}\frac{q_3}{q_1}\frac{d(B/q_1)}{d(q_2/q_1)} - (1 - p_1/q_1)C_1E_1 .$$

Splitting elasticities into income and substitution effects we may combine terms 1 and 4 and 3 and 5 to simplify this expression as before. Using also the fact already noted that substitution elasticities sum to zero, we obtain, corresponding to equation (9) of Part 2:

$$(13)\quad E_1 = \alpha\Big[\frac{X_1}{x_1}\sigma_{11} - \frac{X_1 - x_1}{x_1}S_{11} - \frac{p_1}{q_1}C_1$$

$$- \left(\frac{X_1}{x_1}\sigma_{13} - \frac{X_1 - x_1}{x_1}S_{13}\right)\left(\frac{q_3}{q_2}\frac{d(q_3/q_1)}{d(q_2/q_1)} - 1\right) - \frac{C_1}{x_1}\frac{q_3}{q_1}\frac{d(B/q_1)}{d(q_2/q_1)}\Big].$$

The problem now is to express the rate of change of q_3/q_1 in terms of elasticities so as to be able to eliminate this from (13). To this end we make use of the fact that the change in total demand for X_3 must be equal to the change in total supply. Thus, using the symbol f for the demand function and g for the supply function, we have

$$f_{q_2/q_1} + f_{q_3/q_1}\frac{d(q_3/q_1)}{d(q_2/q_1)} + f_{C/q_1}\frac{d(C/q_1)}{d(q_2/q_1)} = g_{q_2/q_1} + g_{q_3/q_1}\frac{d(q_3/q_1)}{d(q_2/q_1)},$$

which on conversion to elasticities yields

$$\frac{q_2}{q_3}\frac{d(q_3/q_1)}{d(q_2/q_1)} = \frac{S_{22} - \varepsilon_{22} - \dfrac{C_2q_2}{X_2q_2}\dfrac{d(B/q_1)}{d(q_2/q_1)} - \dfrac{C_2q_2}{X_2q_2}(X_2 + x_2) + \dfrac{(q_1 - p_1)}{q_2}\dfrac{x_1}{X_2}C_2E_1}{\varepsilon_{22} - S_{22} + C_2}.$$

Thus

$$\left(\frac{q_2}{q_3}\frac{d(q_3/q_1)}{d(q_2/q_1)} - 1\right) = \frac{\sigma_{21} - S_{21} - \dfrac{q_2x_2}{q_2X_2}C_2 - \dfrac{C_2q_2}{X_2q_2}\dfrac{q(B/q_1)}{d(q_2/q_1)} + \dfrac{(q_1 - p_1)}{q_2}\dfrac{x_1}{X_2}C_2E_1}{\sigma_{22} - S_{22}}. \tag{14}$$

We now define

$$\lambda' = \alpha'/\left\{1 + \frac{q_1 - p_1}{q_2}\left[\frac{\alpha'(X_1\sigma_{12} - (X_1 - x_1)S_{12})}{X_2\sigma_{22} - X_2S_{22}}\right]C_2\right\}.$$

λ' is clearly a multiplier analogous to α', due to the tariff loading. Since $p_1X_1\sigma_{12} + p_2X_2\sigma_{22} + p_3X_3\sigma_{23} = 0$, and in the normal case where there is no complementarity in consumption σ_{12} and σ_{23} are positive, and since supply elasticities have similar properties of opposite sign, the expression inside the square brackets must be negative and numerically less than unity. It follows that λ' must be positive and greater than unity, but by a small amount. It should be noted, however, that if complementarity is present between goods 2 and 3 (i. e., if σ_{23} is negative), σ_{13} may be numerically greater than σ_{33}, and λ' could conceivably be negative and large. We shall discover it to be a general rule that a high degree of complementarity in consumption can lead to perverse results. But this is precisely the situation where equilibrium in international trade might not be possible.

Substituting (14) into (13), making use of our definition of λ', we obtain

$$\begin{aligned} E_1 = \lambda'\Bigg[&\frac{X_1}{x_1}\sigma_{11} - \frac{p_1}{q_1}C_1 - \frac{X_1 - x_1}{x_1}S_{11} \\ &- \left(\frac{X_1}{x_1}\sigma_{12} - \frac{X_1 - x_1}{x_1}S_{12}\right)\left(\sigma_{21} - \frac{x_2}{X_2}\frac{q_2}{q_2}C_2 - S_{21}\right)\Big/(\sigma_{22} - S_{22}) \\ &+ \frac{d(B/q_1)}{d(q_2/q_1)}\frac{q_2}{q_2}\frac{C_2}{X_2}\left(\frac{X_1}{x_1}\sigma_{12} - \frac{X_1 - x_1}{x_1}S_{12}\right)\Big/(\sigma_{22} - S_{22}) \\ &- \frac{C_1}{x_1}\frac{q_2}{q_1}\frac{d(B/q_1)}{d(q_2/q_1)}\Bigg]'. \end{aligned}$$

A similar expression may be obtained for E_2 so that, by substitution in (5), we obtain the final formula[16]

$$\beta = -x_1\left[\frac{U}{V}+1\right], \tag{15}$$

where

$$U = \lambda'\left[\left(\frac{X_1}{x_1}\sigma_{11}-\frac{X_1-x_1}{x_1}S_{11}\right)-\left(\frac{X_1}{x_1}\sigma_{12}-\frac{X_1-x_1}{x_1}S_{12}\right)(\sigma_{21}-S_{21})/(\sigma_{22}-S_{22})\right]'$$
$$+\lambda\left[\left(\frac{X_2}{x_2}\sigma_{21}-\frac{X_2-x_2}{x_2}S_{22}\right)-\left(\frac{X_2}{x_2}\sigma_{22}-\frac{X_2-x_2}{x_2}S_{22}\right)(\sigma_{21}-S_{22})/(\sigma_{22}-S_{22})\right],$$

$$V = 1-\frac{\lambda' p_1}{q_1}C_1'+\frac{\lambda'[q_1X_1\sigma_{12}-q_1(X_1-x_1)S_{12}]'}{(q_2X_2\sigma_{22}-q_2X_2S_{22})'}\frac{p_1C_2'}{q_1}$$
$$-\frac{\lambda q_2}{p_2}C_2+\frac{\lambda[p_2X_2\sigma_{22}-p_2(X_2-x_2)S_{22}]}{(p_2X_2\sigma_{22}-p_2X_2S_{22})}\frac{q_2C_2}{p_2}.$$

Equations (14) and (15) constitute our fundamental results.

4. INTRODUCTION OF MORE REALISTIC ASSUMPTIONS

The formula (15) which we have just derived is not by any means as formidable as it looks. In fact, it is basically the same as the simple two-good result already discussed, having much the same properties of sign and magnitude. This may best be demonstrated by tracing the various modifications required to accommodate successively more realistic assumptions.

[16] It would be wrong to introduce this formula without explaining why it differs from that for the ratio t/v which may be derived from Professor Meade's Table X, (*The Balance of Payments: Mathematical Supplement*, p. 156).

In the process of developing his basic model Professor Meade defines elasticity of supply as the percentage change in output due to a one per cent change in the *reciprocal of the marginal product of labor*, homogenous labor being the only factor of production (*loc. cit.*, p. 8, also pp. 22–23). This definition is used throughout. It is rationalized as follows. The wage rate is always equal to the marginal value product of labor, so that the reciprocal of the marginal product of labor is identical with the ratio of product price to the wage rate. To define supply as a function of the reciprocal of the marginal product of labor, therefore, is equivalent to defining supply as a function of the product price divided by the constant wage rate. Professor Meade's elasticity of supply is the percentage change in output due to a one per cent change in the product price *with the wage rate held constant*. Unfortunately this is not the same as the more usual definition of elasticity, namely, the percentage change in output due to a one per cent change in the product price *with all other product prices held constant*. The theory of production tells us that wage rates are a function of all product prices, so that to change one product price holding all others constant necessarily implies a change in the wage rate. Alternatively, to change any one product price holding wage rates constant necessarily implies that at least one other product price must have changed in compensation.

In the special case where supply elasticities are zero and there is no home production of importables, (15) above and Professor Meade's t/v are identical

Suppose first of all that we have:

(i) No tariffs or transport costs ($p = q$);

(ii) Two classes of goods only, imports and exports;

(iii) No production by any country of the good imported;

(iv) Zero elasticities of supply.

Formula (15) reduces on these assumptions to

$$\beta = -x_1\left[\frac{\sigma'_{11} + \sigma_{22}}{C_1 - C'_1} + 1\right]. \tag{16}$$

This is Professor Meade's result, already commented upon in Part 2. Now consider the effect of introducing a class of non-traded goods. That is, relax Assumption (ii) above but retain Assumptions (i), (iii) and (iv). In this case

$$\beta = -x_1\left[\frac{\left(\sigma_{11} - \frac{\sigma_{13}\sigma_{31}}{\sigma_{33}}\right)' + \left(\sigma_{22} - \frac{\sigma_{23}\sigma_{32}}{\sigma_{33}}\right)}{1 - C'_1 - C_2 + \left(\frac{p_1X_1\sigma_{13}}{p_3X_3\sigma_{33}}\right)'C'_3 + \left(\frac{p_2X_2\sigma_{23}}{p_3X_3\sigma_{33}}\right)C_3} + 1\right]. \tag{17}$$

This is a very straightforward generalization. Consider first the numerator of the first term inside the brackets. We know from the theory of demand that $(\sigma_{11}\sigma_{33} - \sigma_{13}\sigma_{31})$ is positive. This is a condition of stability. As σ_{33} is bound to be negative, it follows that $\sigma_{11} - (\sigma_{13}\sigma_{31}/\sigma_{33})$ must therefore be negative. In fact, this expression has properties analogous to the σ_{11} which it replaces as we move from (16) to (17). It tends to zero when good 1 and good 3 are always consumed in fixed proportions. It tends to infinity when good 1 and good 3 are perfect substitutes. Like σ_{11} in a two-good world this new expression is a measure of the substitution possibilities in the system. With three goods, however, substitutability between goods 1 and 2 will not help in changing the level of foreign demand for good 1 if there is complete complementarity between goods 1 and 3. This is true because the consumption of the non-traded good is fixed by the given supply, so that the consumption of good 1 must also be fixed, whatever the price.

Consider now the denominator of the first term inside the brackets. For any one country we know from the theory of demand that

$$p_1X_1\sigma_{13} + p_2X_2\sigma_{23} + p_3X_3\sigma_{33} = 0\,.$$

Suppose for the moment that both countries have:

(a) Identical substitution elasticities,

(b) Identical average propensities to consume, and

(c) Identical marginal propensities to consume non-traded goods.

Then the complicated expression under examination reduces at once to $C_1 - C_1'$ and corresponds exactly with the two-good result.

It is also clear that however much elasticities differ between countries all terms of the denominator remain less than or equal to one. Its total magnitude must be of the order of $C_1 - C_1'$ in any case. Moreover, the fact that one region is small while the other is the rest of the world does not give us any cause to look for differences in parameters. It is true, of course, that the greater the number of individual consumers the greater will be the absolute change in sales due to a one per cent change in price. But the *percentage* change in sales is not necessarily different with different numbers. It is for this reason that the rather cumbersome elasticity forms have been retained. We emphasize also that the C's in our formulas are marginal propensities to *consume importables* and not marginal propensities to import. Marginal propensities to import will be affected by the relative size of the trading areas, but marginal propensities to consume importables will not.

It is clear that all the arguments put forward in Part 2 regarding the sign and magnitude of β in the two-good case apply with equal force when the extra good is introduced. The numerator of the first term of (17) inside the square brackets consists of the sum of two negative terms which may vary from zero to infinity. The denominator is not very different from the difference between two positive marginal propensities, both of which must be less than one. Moreover, the numerator will be numerically small only when there is high complementarity among all three goods, a circumstance in which the denominator is all the more likely to be even smaller. High complementarity implies that countries will have similar substitution elasticities and similar marginal and average propensities to consume for all goods, which would imply a zero denominator. In sum, it seems highly unlikely that the first term of (17) could ever be numerically less than one, so that the sign of β will depend entirely on the sign of the denominator. It follows that the sign of β is uncertain even when elasticities are high.

We now proceed to show how the uncertainty of sign persists *a fortiori* when all other assumptions are relaxed. To set out in full the equation for β when Assumptions (iii) and (iv) are removed would be tedious. The reader is invited therefore to turn back to (15) in Part 3 and to read all λ's as unity and all q's as p's. It will be seen that once again the essential form of (16) and (17) is preserved. To reduce the denominator V of the first term inside the square brackets

to $C_1 - C_1'$ as before, we need in addition to those already listed the extra assumptions:

(a) That elasticities of supply are the same in both countries, and

(b) That the *average* propensity to produce every good is the same in both countries, the propensity to produce being the exact supply analogue of the average propensity to consume.

All of those conditions will, of course, not be met. On the other hand, there is no reason to expect the bias to go one way rather than another. Reasons for this on the demand side have already been given. On the supply side, it would be natural to suppose that in any country the average propensity to produce importables would be lower than in the region from which it imports. But for this very reason supply elasticity is likely to be greater; it is easier to achieve a ten per cent increase in production when absolute production is low than when it is high. Inspection of our formula reveals that these two biases work in opposite directions. The sign of our denominator remains indeterminate, and its magnitude is still of the order of the difference between the two positive marginal propensities to consume good 1.

Looking at the new numerator we notice that this takes exactly the same form as in (17), except that an elasticity of supply is added to each substitution elasticity. In addition, multipliers equal to the ratio of consumption of importables to imports, similar to those encountered in Part 2, reappear. Making use of the fact that stability conditions require both $S_{11}S_{33} - S_{13}S_{31}$ and $\sigma_{11}\sigma_{33} - \sigma_{13}\sigma_{31}$ to be positive, we can easily prove that the new numerator, like the old, must be negative. We have

$$(\sigma_{11}S_{33})(\sigma_{33}S_{11}) \geqq (S_{13}\sigma_{31})(S_{31}\sigma_{13}) .$$

By the symmetry of substitution effects

$$(X_1 - x_1)S_{13}/p_3 = X_3S_{31}/p_1 \text{ and } X_1\sigma_{13}/p_3 = X_3\sigma_{31}/p_1 ,$$

so that

$$(X_1 - x_1)S_{13}\sigma_{31} = X_1S_{31}\sigma_{13} .$$

It is well known that the geometric mean of two positive numbers must be less than or equal to their arithmetic mean, and that the geometric mean is equal to the arithmetic mean when the two numbers are equal; thus

$$\frac{X_1}{x_1}\sigma_{11}S_{33} + \frac{X_1 - x_1}{x_1}S_{11}\sigma_{33} \geqq 2\left[\left(\frac{X_1}{x_1}\sigma_{11}S_{33}\right)\left(\frac{X_1 - x_1}{x_1}S_{11}\sigma_{33}\right)\right]^{\frac{1}{2}}$$

$$\geqq 2\left[\left(\frac{X_1}{x_1}S_{31}\sigma_{13}\right)\left(\frac{X_1 - x_1}{x_1}S_{13}\sigma_{31}\right)\right]^{\frac{1}{2}} = \frac{X_1 - x_1}{x_1}S_{13}\sigma_{31} + \frac{X_1}{x_1}S_{31}\sigma_{13} .$$

Given that $(\sigma_{33} - S_{33})$ is negative, it follows at once that

$$\left(\frac{X_1}{x_1}\sigma_{11} - \frac{X_1 - x_1}{x_1}S_{11}\right) - \left(\frac{X_1}{x_1}\sigma_{13} - \frac{X_1 - x_1}{x_1}S_{13}\right)(\sigma_{31} - S_{31})/(\sigma_{33} - S_{33})$$

is negative or zero.

It can also be shown that the whole expression tends to zero if and only if *both*

(a) Goods 1 and 3 are always demanded in fixed proportions, *and*
(b) Goods 1 and 3 are in joint supply, that is can be supplied only in fixed proportions.

On the other hand, it tends to infinity if *either*

(a) Goods 1 and 3 are perfect substitutes, *or*
(b) Supply is perfectly elastic in the sense that relative costs are unchanged when resources are shifted from production of good 1 to good 3 or vice versa.

All of this means that the introduction of supply elasticities tends to increase the magnitude of the first term inside the square brackets in the three-good just as in the two-good case. The sign of β will be determined normally by the sign of this term, which is determined in turn by the sign of the denominator V. Perverse cases are impossible except when there is both complementarity in consumption and production among all three goods, and this, as we have seen, is just the case where a zero denominator is probable. All of the arguments used in the two-good case and the case of three goods with zero supply elasticities apply generally with greater force, partly because flexibility in supply tends to make the numerator U larger and partly because elasticities have positive multipliers X/x which cannot be less than unity and may be quite large. In particular, one of these multipliers will be large when one trading region is small relative to the other.

Finally, the admission of tariff barriers causes the positive multipliers λ to appear as shown in formula (15). As shown also in Part 3 these mulitipliers will normally be positive and slightly greater than unity. This still further increases the magnitude of the numerator U of the first term inside the square brackets without significantly changing the magnitude of the denominator V. In fact, in the denominator, λ wherever it appears is multiplied by p_1/q_1 or q_2/p_2, both of which are positive and slightly less than unity. The arguments above therefore are further reinforced by the existence of tariffs. We have to conclude that an improvement in the balance of trade may require a movement of the real terms of trade either way even though elastic-

ities may be high.

We have now to ask what change in the relative price of non-traded goods is necessary. Here it is possible to be a little more definite. By a method analogous to that used to obtain equation (14) in Part 3 we may derive an expression for the percentage change in the ratio of the prices of non-traded goods and exportables which must be associated with a one per cent adverse movement of the real terms of trade if full employment is to be maintained. We write this

$$(18) \qquad (S_{31} - \sigma_{32})/(\sigma_{33} - S_{33}) + \left(\frac{p_2 x_2}{p_3 X_3} + \frac{p_1 x_1}{p_3 X_3}\frac{\beta}{x_1}\right) C_3/(\sigma_{33} - S_{33}) .$$

For convenience the term which arises because of the existence of tariffs has been left out on the ground that it will in any case be small relative to β (cf. equation (14), Part 3).

Consider first the case where a one per cent adverse movement of the real terms of trade improves the trade balance (β positive). The entire argument above suggests that β/x_1 will be large relative to $(\sigma_{33} - S_{33})$; the denominator in β is a small number not very different from $C_1 - C_1'$. We have already noted that unless there is complementarity in consumption or joint supply between goods 1 and 3, $(S_{31} - \sigma_{32})/(\sigma_{33} - S_{33})$ will be numerically less than one. Thus, the sign of the whole expression above is likely to be dominated by the sign of $\beta/(\sigma_{33} - S_{33})$. If the balance of trade is improved (β is positive), the sign of the whole will be negative. The price p_3 must fall relative to the price p_1.

To find the change in p_3/p_1 when a one per cent *favorable* movement in the terms of trade is required to improve the trade balance, the expression above must be multiplied by (-1), and we must keep in mind that this time β is negative. $-(S_{31} - \sigma_{32})/(\sigma_{33} - S_{33})$ is now negative. On the other hand, $-(p_2 x_2 C_3/p_3 X_3(\sigma_{33} - S_{33}))$ is positive but will usually be considerably less than one. By the argument of the preceding paragraph, therefore, p_3 falls relative to p_1 in this case also. Thus we have the general conclusion:

(i) That normally the price of non-traded goods must fall relative to the price of exportables if the balance of trade is to be improved.

This conclusion is not inescapable. It might be reversed, but only if there were high complementarity between exportables and non-traded goods at home and similarly abroad, as well as low elasticities of supply. β might be small, even though $(\sigma_{33} - S_{33})$ is not, and $(S_{31} - \sigma_{32})/(\sigma_{33} - S_{33})$ might be positive and greater than one.

Next it is obvious that if p_3/p_1 has to rise to improve the trade balance and p_3/p_1 has to fall, then p_3/p_2 must fall also. If, on the other hand, p_2/p_1 falls by one per cent and the trade balance improves, then p_3/p_2 will also fall only if the fall in p_3/p_1 is greater than one per cent. In view, however, of what has been said about the magnitude of β and hence of (18) this seems certain to be the case. Thus we shall argue:

(ii) That normally the price of non-traded goods must fall relative to the price of importables if the balance of trade is to be improved.

We note again that the existence in the system of complementarity and/or joint supply may upset this conclusion. In view of the fact that our "goods" are really classes of goods of many different kinds neither complementarity nor joint supply is likely to be encountered in practice. In any event, any case where they do occur will be immediately recognizable.

Conclusions and remarks. We conclude that an improvement in the balance of trade requires, in order of importance:

(1) A cut in money spending equal to
  (a) the trade balance improvement, plus
  (b) the money equivalent of the real gain or loss due to the change in the terms of trade (cf. (3)), plus
  (c) the money equivalent of any change in tariff revenue due to the change in the terms of trade.
(2) A fall in the price of non-traded goods relative to that of traded goods.
(3) Some change in the real terms of trade which may be positive or negative. *This will be small relative to change* (2).

Conclusion (2) may be upset by the existence of complementarity and/or joint supply.

Of these conclusions, (3) accords least with our intuition. The existence of a trade deficit implies that the country in deficit is living beyond its means, and that excess spending is financed by a loss of reserves. It seems a little odd that even in quite ordinary cases an improvement in the real terms of trade which increases real income might be required to correct a disorder due to too high expenditure. The explanation of the paradox is simple. Early preoccupation with the role of fluctuating exchange rates in adjusting the supply and demand for currency has encouraged economists to look on depreciation as the direct cause of any improved trade balance. Moreover, we have been conditioned by two-good models to think of depreciation

and an adverse movement of the real terms of trade as synonymous. Thus it is felt that only devaluation can make a country worse off, as it ought to be if its trade balance is to be improved. In fact, however, it is primarily the essential cut in money spending which corrects the trade deficit and not the change in the real terms of trade. And it is the cut in expenditure which brings about the required reduction in real wealth.

The common sense of this conclusion becomes evident as soon as we reflect that there could be a case where no relative price changes are required at all. In the simple two-good model a cut in home spending might release goods (both importables and exportables) in exactly the proportions which would be demanded abroad following increased spending there. If the changes in spending were made equal to the trade deficit, this would be corrected without any need to change relative prices. In terms of our model, $C_1 = C_1'$ so that β is infinite and a negligible change in p_2/p_1 will correct any finite imbalance. Put this way we see that if zero change in p_2/p_1 is not "perverse," then a situation requiring an improvement (negative change in p_2/p_1) should not be any more surprising than the opposite.

The case where no change in relative prices is called for reveals the element of truth in the argument that deflation by itself can cure a trade deficit. But it must be recognized that this can never be wholly true when there exists a class of non-traded goods. A cut in spending will cut the demand for non-traded goods which cannot be sold abroad. To maintain full exployment it is absolutely necessary to reduce the price of non-traded goods relatively.

This brings us to the common sense of condition (2) above. The price of non-traded goods has to fall relative to the prices of exportables and importables simply because the effect on their sale of the cut in home spending is not offset by any increase in spending abroad. The fall in home demand for exportables and importables due to the cut in home spending is, on the other hand, partly or even wholly cancelled out by increased demand abroad. This is most clearly seen in the case outlined above where the relative price of exportables and importables does not change. But similar considerations are always present.

Finally, we come to the question of exchange depreciation as a policy. We have remarked in Part 1 that the purpose of this paper is to define ends rather than means, and indeed no certain policy can be defined except in terms of a dynamic model. One thing, however, is certain. The impact effect of exchange depreciation will be to

Next it is obvious that if p_3/p_1 has to rise to improve the trade balance and p_3/p_1 has to fall, then p_3/p_2 must fall also. If, on the other hand, p_2/p_1 falls by one per cent and the trade balance improves, then p_3/p_2 will also fall only if the fall in p_3/p_1 is greater than one per cent. In view, however, of what has been said about the magnitude of β and hence of (18) this seems certain to be the case. Thus we shall argue:

(ii) That normally the price of non-traded goods must fall relative to the price of importables if the balance of trade is to be improved.

We note again that the existence in the system of complementarity and/or joint supply may upset this conclusion. In view of the fact that our "goods" are really classes of goods of many different kinds neither complementarity nor joint supply is likely to be encountered in practice. In any event, any case where they do occur will be immediately recognizable.

Conclusions and remarks. We conclude that an improvement in the balance of trade requires, in order of importance:

(1) A cut in money spending equal to
  (a) the trade balance improvement, plus
  (b) the money equivalent of the real gain or loss due to the change in the terms of trade (cf. (3)), plus
  (c) the money equivalent of any change in tariff revenue due to the change in the terms of trade.

(2) A fall in the price of non-traded goods relative to that of traded goods.

(3) Some change in the real terms of trade which may be positive or negative. *This will be small relative to change* (2).

Conclusion (2) may be upset by the existence of complementarity and/or joint supply.

Of these conclusions, (3) accords least with our intuition. The existence of a trade deficit implies that the country in deficit is living beyond its means, and that excess spending is financed by a loss of reserves. It seems a little odd that even in quite ordinary cases an improvement in the real terms of trade which increases real income might be required to correct a disorder due to too high expenditure. The explanation of the paradox is simple. Early preoccupation with the role of fluctuating exchange rates in adjusting the supply and demand for currency has encouraged economists to look on depreciation as the direct cause of any improved trade balance. Moreover, we have been conditioned by two-good models to think of depreciation

and an adverse movement of the real terms of trade as synonymous. Thus it is felt that only devaluation can make a country worse off, as it ought to be if its trade balance is to be improved. In fact, however, it is primarily the essential cut in money spending which corrects the trade deficit and not the change in the real terms of trade. And it is the cut in expenditure which brings about the required reduction in real wealth.

The common sense of this conclusion becomes evident as soon as we reflect that there could be a case where no relative price changes are required at all. In the simple two-good model a cut in home spending might release goods (both importables and exportables) in exactly the proportions which would be demanded abroad following increased spending there. If the changes in spending were made equal to the trade deficit, this would be corrected without any need to change relative prices. In terms of our model, $C_1 = C_1'$ so that β is infinite and a negligible change in p_2/p_1 will correct any finite imbalance. Put this way we see that if zero change in p_2/p_1 is not "perverse," then a situation requiring an improvement (negative change in p_2/p_1) should not be any more surprising than the opposite.

The case where no change in relative prices is called for reveals the element of truth in the argument that deflation by itself can cure a trade deficit. But it must be recognized that this can never be wholly true when there exists a class of non-traded goods. A cut in spending will cut the demand for non-traded goods which cannot be sold abroad. To maintain full exployment it is absolutely necessary to reduce the price of non-traded goods relatively.

This brings us to the common sense of condition (2) above. The price of non-traded goods has to fall relative to the prices of exportables and importables simply because the effect on their sale of the cut in home spending is not offset by any increase in spending abroad. The fall in home demand for exportables and importables due to the cut in home spending is, on the other hand, partly or even wholly cancelled out by increased demand abroad. This is most clearly seen in the case outlined above where the relative price of exportables and importables does not change. But similar considerations are always present.

Finally, we come to the question of exchange depreciation as a policy. We have remarked in Part 1 that the purpose of this paper is to define ends rather than means, and indeed no certain policy can be defined except in terms of a dynamic model. One thing, however, is certain. The impact effect of exchange depreciation will be to

raise the price of importables relative to non-traded goods. Moreover, since they can be sold abroad at a higher price, the pressure will be towards a higher home price for exportables as well. *It may be that the success of exchange depreciation as a policy rests more upon its power to reduce the price of non-traded goods relative to those traded than upon its power to affect the real terms of trade.*

Australian National University, Australia

INTERNATIONAL ECONOMIC REVIEW
Vol. 2, No. 1, January, 1961

EFFICIENT INTERNATIONAL TRADE AND EQUALIZATION OF FACTOR PRICES*

BY STANLEY REITER[1]

1. INTRODUCTION

THIS PAPER offers contributions to two sections of international trade theory. First, it extends the analysis of international specialization of production in several ways. It considers the case of both non-linear and linear technologies, possibly differing among countries, and mobile and non-mobile products and resources. It also shows the relation between nationally efficient and internationally efficient production patterns, without prescribing the institutional framework of production. On the basis of this analysis we prove the welfare theorem of classical international trade theory, namely, that the equilibria of the system of free trade and domestic competition are internationally and domestically efficient patterns of production, thus extending the theorems of earlier international trade theory. Second, we use this more general model to derive both necessary and sufficient conditions for factor price equalization, rather than merely sufficient conditions, as has been the case hitherto. We use these conditions to analyze the logical role of the sufficient conditions employed by Professor Samuelson in his discussion of factor price equalization. We do not assume that all countries have identical technologies, or that constant returns to scale prevail; we show that the condition interpreted by Samuelson as "non-reversibility of relative factor intensities" need not be imposed on the production functions, although an analogous condition is used elsewhere. Our argument is presented for many countries, many products, and many factors. The existence of by-products is recognized, but joint production is excluded. Transportation appears explicitly in the model. The analysis is independent of the international distribution of resources and of the conditions of demand.

* Manuscript received December 30, 1959, revised August 29, 1960.

[1] I wish to express my gratitude to Professor Leonid Hurwicz of the University of Minnesota for his assistance. I am also indebted to Professor Tjalling C. Koopmans of Yale University and to Professor Kenneth J. Arrow of Stanford University for their comments on this paper. Of course they are not to be held responsible for errors or omissions. The results given here are contained in Technical Report No. 13, under Contract N6onr-25133 (NR-047-004), Department of Economics, Stanford University, August 16, 1954. The author acknowledges the support of the Office of Naval Research under Contract Nonr 1100 (16) (NR-047-026).

2. EFFICIENT INTERNATIONAL LOCALIZATIONS OF PRODUCTION

2.1. *Technology.* Assume that the world technology and the technologies of the constituent countries are of the Kuhn-Tucker type; that is, they are represented by non-linear activity analysis models.[2]

Let U denote a finite dimensional linear space, the world commodity space. The elements $u \in U$ are commodity vectors; such a vector denotes a specification of the quantity (flow) of each commodity in each country.

Let X denote the non-negative part of a finite dimensional linear space; X is the activity space. The elements $x \in X$ are activity vectors, or vectors of activity levels.

The world technology is given by a mapping

$$L : X \to U .$$

The image $L(x) \in U$ of a point $x \in X$ is the vector of inputs and outputs associated with the activity vector x. Thus, if production is carried out in every country at the levels x, then the resulting commodity flows are $u = L(x)$.

The world technology may be decomposed in terms of the constituent technologies of the various countries.

Let ${}^{(j)}U$, $j = 1, 2, \cdots, J$, denote a finite dimensional linear space called the commodity space of country j. The elements ${}^{(j)}u \in {}^{(j)}U$ are commodity vectors denoting a specification of the quantity of each commodity in country j. We have, then, the world commodity space as the Cartesian product of the commodity spaces of the component countries[3]

$$U = {}^{(1)}U \otimes \cdots \otimes {}^{(J)}U .$$

Similarly the world activity space may be decomposed to show its structure in terms of the activity spaces of the individual countries. Here, however, it is desirable to distinguish the transportation activities which connect the various national technologies.

A transportation activity has the effect of transporting a quantity of some specified commodity from one country to another. The level of the activity determines the quantity transported. The activity is thought of as being part of the technology of the exporting country

[2] H. W. Kuhn and A. W. Tucker, "Non-linear Programming," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman (Berkeley: University of California Press, 1951), pp. 481-92.

[3] Let A and B be sets, then by $A \otimes B$ we mean the set of all pairs (a, b) with $a \in A$ and $b \in B$. $A \otimes B$ is called the Cartesian product of A and B. The Cartesian product of any number of sets is defined similarly

and hence its level is a coordinate of the activity space of that country. We assume that transportation is free in terms of resources and that if a commodity can be transported between any pair of countries, then it can be transported between every pair.

Write the world activity space X as the Cartesian product of two spaces:

$$X = \tilde{X} \otimes T ,$$

where $\tilde{X}$ is the activity space of other than transportation activities and T is the space of transportation activities. Further, $\tilde{X}$ and T have the structure

$$\tilde{X} = {}^{(1)}\tilde{X} \otimes \cdots \otimes {}^{(J)}\tilde{X} , \qquad T = {}^{(1)}T \otimes \cdots \otimes {}^{(J)}T ,$$

where ${}^{(j)}\tilde{X}$ is the space of activities other than transportation in the j-th country and ${}^{(j)}T$ is the space of transportation activities in country j.

It is useful to consider activity vectors of the individual countries as elements of the world activity space. The following notational convention permits this. Putting the prescript j before an element of the world activity space will equate to zero all components which do not refer to the j-th country.

Having factored the activity space, we must now decompose the mapping L in a corresponding way.

The technology of country j is given by the restriction of L to ${}^{j}X$.[4] For convenience of notation, write

$${}^{j}L = L_{{}^{j}X} \qquad \text{for } j = 1, 2, \cdots, J .$$

Then we may write

$$L(x) = \sum_{j=1}^{J} {}^{j}L({}^{j}x) \qquad \text{for all } x \in X .$$

Thus, ${}^{j}L({}^{j}x) \in U$ is the vector of commodity inputs and outputs when the various available activities (including transportation) are carried out at the levels ${}^{j}x$ in country j.

Having already defined the mapping L and ${}^{j}L$, we need only consider their restrictions to $0 \otimes T \subset X$ and ${}^{j}T \subset X$ to isolate the transportation activities. A typical vector $t \in T$ has the form

$$t = (\cdots, {}^{k}t_i^s, \cdots) ,$$

[4] Let A and B be sets with $A \subset B$. Let f be a function (mapping) defined on B. We define the function f_A on A by

$$f_A(x) \equiv f(x) \qquad \text{for all } x \in A$$

and call f_A the restriction of f to A. Loosely speaking, f_A is the function f considered only on A.

and the restriction $L_{0\otimes T}$ of L to $0 \otimes T$ has the form

$$L_{0\otimes T} = \sum_{i,k,\delta} {}^{i}t_{\delta}^{k}({}^{k}d_{\delta} - {}^{i}d_{\delta}) ,$$

where

$${}^{k}d_{\delta} = (0, \cdots, {}^{k-1}0, {}^{k}d_{\delta}, {}^{k+1}0, \cdots, 0) .$$

Here i and k range over the list of countries, while δ ranges over a specified subset of the list of commodities, namely, those for which there is a transportation activity available. Thus, if the commodity δ is transportable, then ${}^{i}t_{\delta}^{k}$ denotes the amount of it which is transported from country i to country k. Notice that for fixed k and δ,

$$\sum_{i} {}^{i}t_{\delta}^{k}({}^{k}d_{\delta})$$

has as the δ-th component to the k-th subvector the total quantity of δ imported by country k. Similarly,

$$\sum_{k} {}^{i}t_{\delta}^{k}(-{}^{k}d_{\delta})$$

has as δ-th component to the i-th subvector the total amount of commodity δ exported to all other countries by country i.

Notice that for every δ,

$$\sum_{k}\sum_{i} {}^{i}t_{\delta}^{k}({}^{k}d_{\delta}) + \sum_{i}\sum_{k} {}^{i}t_{\delta}^{k}(-{}^{i}d_{\delta}) = 0 ,$$

which is to say that the net world export of every commodity is, as it should be, zero.

A commodity for which transportation activities are available will be called internationally mobile, or briefly, mobile; those commodities for which transportation activities are not available will be called immobile.

One of the essential features of the model to be used here is the distinction between products and resources. This distinction will now be drawn. Suppose that the list of commodities can be broken down into two sub-lists, one a list of products, the other a list of resources or factors of production.

For each j, factor the commodity space ${}^{j}U$ of country j into two component spaces, ${}^{j}Y$ and ${}^{j}Z$, where ${}^{j}Y$ is the space of products and ${}^{j}Z$ is the space whose coordinates refer to resources. Thus, ${}^{j}U$ is the Cartesian product of ${}^{j}Y$ and ${}^{j}Z$, i.e.,

$${}^{j}U = {}^{j}Y \otimes {}^{j}Z .$$

Elements ${}^{j}u \in {}^{j}U$ may then be written

$$^{(j)}u = \begin{pmatrix} ^{(j)}y \\ ^{(j)}z \end{pmatrix}, \qquad ^{(j)}y \in {}^{(j)}Y, \qquad ^{(j)}z \in {}^{(j)}Z.$$

We assume that this classification is the same throughout the world and hence write

$$U = Y \otimes Z,$$

where

$$Y = {}^{(1)}Y \otimes \cdots \otimes {}^{(J)}Y, \qquad Z = {}^{(1)}Z \otimes \cdots \otimes {}^{(J)}Z.$$

Correspondingly, the coordinates of the mapping L may be relabelled. Let

$$L(x) = \begin{pmatrix} G(x) \\ H(x) \end{pmatrix} \qquad x \in X,$$

where

$$G\colon X \to Y, \qquad H\colon X \to Z,$$

and define

$$^{j}G\colon X \to Y, \qquad ^{j}H\colon X \to Z,$$

to be the restrictions of G and H respectively to ^{j}X for each $j = 1, 2, \cdots, J$.

In keeping with the notation above, the transportation activities can also be separated into those operating on products as against those operating on resources. Thus, write G_T and H_T for the restrictions of G and H respectively to $0 \otimes T$, and, correspondingly, $^{j}G_{k_T}$ and $^{j}H_{k_T}$ for the restrictions of ^{j}G and ^{j}H to $0 \otimes {}^{k}T$. Similarly, $\tilde{G}$ and $\tilde{H}$, and $^{j}\tilde{G}$, $^{j}\tilde{H}$, will denote the restrictions of G and H to $\tilde{X} \otimes 0$ and $^{j}\tilde{X} \otimes 0$ respectively. Thus $^{j}\tilde{G}$ denotes the flow of products in country j when there is no transportation.

Each country j is assumed to be endowed with some fixed quantity (possibly zero) of each resource. The resource endowment of the world is given by a vector

$$c \in Z.$$

The J sub-vectors

$$^{j}c \in {}^{j}Z \qquad (j = 1, 2, \cdots, J)$$

into which c can be factored express the resource endowments of each of the J countries in turn.

The limited availability of resources imposes a restriction on production, namely, that the activities employed may not require inputs in excess of the amounts available, including imports. This restriction

is expressed by the constraints

$$ {}^{j}H({}^{j}x) \leqq {}^{j}c + \sum_{i \neq j} {}^{i}H_{ij}({}^{i}t^{j}) \qquad (i, j = 1, 2, \cdots, J) . $$

Notice that if no transportation of resources is possible, these inequalities express the usual restriction imposed on production in each country by the requirement that its use of inputs not exceed its resource endowment.

Equivalently, we may define

$$ {}^{j}F({}^{j}x) \equiv {}^{j}c + \sum_{i \neq j} {}^{i}H_{ij}({}^{i}t^{j}) - {}^{j}H({}^{j}x) \qquad (i, j = 1, 2, \cdots, J) , $$

and require

$$ {}^{j}F({}^{j}x) \geqq 0 \qquad (j = 1, 2, \cdots, J) . $$

From the viewpoint of the world as a whole, define

$$ F(x) = \begin{bmatrix} {}^{1}F({}^{1}x) \\ {}^{2}F({}^{2}x) \\ \vdots \\ {}^{J}F({}^{J}x) \end{bmatrix} . $$

The requirement is, then,

$$ F(x) \geqq 0 . $$

In the preceding discussion two classifications of commodities have been made, resulting in the four categories of commodities shown in the table below:

	Product	Resource
Internationally mobile	I	III
Internationally immobile	II	IV

The classical theorists assumed that Classes II and III are empty, i.e., that all products are internationally mobile, and all resources internationally immobile. Indeed, some have taken this to be part of the definition of international trade.[5] We assume that each of the four classes may have members and extend the notation as follows in order to facilitate drawing the various distinctions: We use the symbol * to indicate a mobile commodity and the symbol ♯ to indicate an immobile one and suppose that the list of commodities is the same for all countries. Then let Class I be indexed by $\alpha^* = 1^*, \cdots, A^*$; Class II

[5] J. Viner, *Studies in the Theory of International Trade* (New York: Harper & Brothers, 1937), p. 597

by $\alpha^1 = 1^1, \cdots, A^1$; Class III by $\beta^* = 1^*, \cdots, B^*$; Class IV by $\beta^1 = 1^1, \cdots, B^1$, where $A^* + A^1 = A$ is the number of products, and $B^* + B^1 = B$ is the number of resources.[6]

This indexing of Classes I through IV corresponds to a listing of the coordinates of the following sub-spaces.

We already have factored the commodity space U as

$$U \equiv Y \otimes Z .$$

Write

$$Y \equiv Y^* \otimes Y^1 , \quad \text{and} \quad Z \equiv Z^* \otimes Z^1 ,$$

where Y^* is the space of mobile products, Y^1 is the space of immobile products, Z^* is the space of mobile resources, and Z^1 is the space of immobile resources. Under the assumption that mobility of commodities is independent of the countries involved, we may write,

$$Y^* \equiv {}^1Y^* \otimes \cdots \otimes {}^JY^* , \qquad Y^1 \equiv {}^1Y^1 \otimes \cdots \otimes {}^JY^1 ,$$
$$Z^* \equiv {}^1Z^* \otimes \cdots \otimes {}^JZ^* , \qquad Z^1 \equiv {}^1Z^1 \otimes \cdots \otimes {}^JZ^1 .$$

2.2 *Efficiency.* Efficiency refers to points in the commodity space; *a point is said to be efficient if it can be produced with the available technology and resources, and if no other point can be produced which gives a larger output of some commodities with no smaller output of any commodity.*

We consider the usual preference or efficiency ordering, $\geq$, defined on the commodity space, where, as usual,

$$a \geq b \text{ means } a_i \geqq b_i \text{ for every } i, \text{ and } a_i \neq b_i \text{ for some } i.$$

Consider the preference ordering as defined only over the space Y of products, rather than extended over the space U of all commodities. There is a slight gain in generality in this procedure, since, by including suitable identity mappings in the technology we may relabel those resources which are in themselves objects of utility so that they appear under an alias in the list of products as well as in the list of resources.

We apply the concept of efficiency to both the commodity space U of the world and separately to the spaces ${}^{(i)}U$ of the individual countries. The concept of efficiency gives rise in a natural way to the related notion of efficiency in the activity space. A point in the activity space is called efficient if its image is an efficient point in the commodity

[6] There is a certain ambiguity in this notation in that it is impossible to distinguish e.g., between $\alpha^* = 3^*$ and $\beta^* = 3^*$. However, since this question will never actually arise, I shall use the notation.

space. Thus,

DEFINITION 2.2.1. *The point $x \in X$ is called world efficient if,*

(1) $$F(x) \geqq 0 ,$$

and

(2) *there is no $x' \in X$ satisfying* (1) *for which $G(x') \geq G(x)$.*

DEFINITION 2.2.2. *The point ${}^{(j)}\tilde{x} \in {}^{(j)}\tilde{X}$ is called locally efficient for $j = 1, 2, \cdots, J$, if,*

(1) $$ {}^{j}F({}^{j}\tilde{x}) \geqq 0 ,$$

and

(2) *there is no ${}^{j}x' \in {}^{j}x$ satisfying* (1) *for which ${}^{j}\tilde{G}({}^{j}\tilde{x}') \geq {}^{j}\tilde{G}({}^{j}\tilde{x})$.*

It is clear that world efficiency and local efficiency are closely related concepts. It follows from the definitions alone that world efficiency of production implies local efficiency of production in every country. The converse, that local efficiency of production in every country implies world efficiency, is not true in general. For the present, without introducing new concepts, it may be said that when no transportation is possible, local efficiency in every country implies world efficiency. These rather obvious facts are stated "for the record" in the following theorem.

THEOREM 2.2.1.[7] (a) *Let $x \in X$ have sub-vectors ${}^{(j)}x \in {}^{(j)}X$, $j = 1, 2, \cdots, J$ as "components." If x is world efficient, then ${}^{(j)}x$ is locally efficient for each $j = 1, 2, \cdots, J$.*

(b) *Let $A^* = B^* = 0$, i.e., let there be no transportation activity in the technology. Let $x \in X$ have "components" ${}^{(j)}\tilde{x} \in {}^{(j)}\tilde{X}$, $j = 1, 2, \cdots, J$. If ${}^{(j)}\tilde{x}$ is locally efficient for each $j = 1, 2, \cdots, J$, then x is world efficient.*

2.3. *Efficiency and prices.* In order to specify more precisely the relationship between world and locally efficient points it is necessary to introduce prices. As many have already observed, prices arise quite naturally in this context, for the concept of efficiency is intimately related to that of intrinsic prices and indeed implies their existence (in the mathematical sense) independent of the existence of such institutions as markets.

The existence of prices associated with efficient points, i.e., efficiency prices, has been established by Koopmans for the case where G and H are linear transformations.[8] A generalization of some of Koopmans'

[7] For a proof of this and subsequent theorems given without proof see S. Reiter, *Efficiency and Prices in the Theory of an International Economy*, University of Chicago Libraries, 1955.

[8] T. C. Koopmans, "Analysis of Production as an Efficient Combination of Activities," *Activity Analysis of Production and Allocation* (New York: John Wiley and Sons, 1951), pp. 33-97.

results to the case where G is a concave function,[9] H is convex, and both are differentiable has been given by Kuhn and Tucker.[10] We shall make extensive use of the Kuhn-Tucker results below and therefore quote the main theorems here.

Kuhn-Tucker Theorems.

DEFINITION 2.3.1. (a) *Let G and F be concave[11] differentiable mappings*

$$G\colon X \to Y\,, \qquad F\colon X \to Z\,.$$

Then $x^0 \in X$ is called a solution of the vector maximum problem if $F(x^0) \geqq 0$ and there is no $x' \in X$ satisfying $F(x') \geqq 0$ with $G(x') \geq G(x^0)$.

The fundamental results of Kuhn and Tucker can now be stated.

THEOREM 2.3.1. (*Equivalence theorem, Theorem 6 of Kuhn and Tucker.*) *Let the mappings*

$$G\colon X \to Y\,, \qquad F\colon X \to Z$$

be concave and differentiable. Then x^0 is a proper[12] solution of the vector maximum problem if and only if for some vector $p > 0$, x^0 and some vector $q^0 \geqq 0$ satisfy

$$pG(x) + q^0F(x) \leqq pG(x^0) + q^0F(x^0) \leqq pG(x^0) + qF(x^0)$$
$$\text{for all } x \geqq 0,\ q \geqq 0.$$

Thus there is a pair of vectors associated with every efficient point. The vector p has the interpretation that its components are prices of those commodities which are desired. These prices must be strictly positive. The components of the vector q are the prices of resources. These prices may be zero but not negative.

Conversely, if for a given positive price vector p there is a non-negative vector q^0 of prices of resources, such that some x^0 satisfies the saddle point inequality of Theorem 2.3.1, then x^0 is efficient.

We also quote the necessary and sufficient conditions for a saddle value of the function

[9] See footnote to Definition 2.3.1.

[10] Kuhn and Tucker, *op. cit.*

[11] Let G be a vector valued function defined on a convex set X. Then G is called convex if

$$G(\theta x_1 + (1-\theta)x_2) \leqq \theta G(x_1) + (1-\theta)G(x_2) \qquad (0 \leqq \theta \leqq 1)\,,$$

i.e., if the line segment joining any two values of the function lies above the graph of the function.

The function G is called concave if $-G$ is convex.

[12] A proper solution is one that can be recognized by differential conditions. See Kuhn and Tucker, *ibid.* for a precise definition.

$$\phi(x, q) \equiv pG(x) + qF(x),$$

which give further insight into the ecomomic relevance of Theorem 2.3.1 and the interpretation of p and q as price vectors.

THEOREM 2.3.2. *Let the mappings*

$$G: X \to Y, \qquad F: X \to Z$$

be concave and differentiable, and for $p > 0$, *let* $\phi(x, q) \equiv pG(x) + qF(x)$. *Then, necessary and sufficient conditions that*

$$\phi(x, q^0) \leqq \phi(x^0, q^0) \leqq \phi(x^0, q) \qquad \text{for all } x \geqq 0,\ q \geqq 0$$

are

$$(1) \qquad \phi_x^0 \leqq 0, \quad \phi_x^0 x^0 = 0, \quad x^0 \geqq 0;$$
$$(2) \qquad \phi_q^0 \geqq 0, \quad \phi_q^0 q^0 = 0, \quad q^0 \geqq 0;$$

where

$$\phi_x^0 = \left[\frac{\partial \phi}{\partial x_i}\right]_{(x^0, q^0)}, \qquad \phi_q^0 = \left[\frac{\partial \phi}{\partial q_k}\right]_{(x^0, q^0)}.$$

These conditions may be interpreted as equilibrium conditions characterizing an organization of production in which each coordinate of x is controlled by a manager, while prices are generated or promulgated by some independent agency.

Condition 1 states that no activity shall earn positive profits as computed using the announced prices, and that no activity which earns negative profits shall be carried out.

Condition 2 states that the resource restriction must not be violated, and that prices of resources in excess availability shall be zero.

2.4. *World and local efficiency and prices in the presence of transportation activities.* In this section we apply the theorems of section 2.3 to make precise the relationship between world and local efficiency when transportation activities exist. As one would expect, if world production is world efficient, then it follows that every component part, so to speak, cannot be improved locally and hence is locally efficient. On the other hand, the local efficiency of production in every country does not by itself insure that world production is world efficient, for this fact gives no assurance that the various economies mesh properly. If all the countries have the same valuation of the commodities which connect them, however, then local efficiency implies world efficiency. This is made precise below.

We use the letter p to denote the world vector of prices of final commodities; the subvector ${}^jp$, $j = 1, \cdots, J$, will denote the prices of

final commodities in the j-th country. Similarly, denote by q the world vector of resource prices and by ${}^jq$, the prices in the j-th country. Recalling the distinction between mobile and immobile commodities, we write ${}^jp_*$ to denote the subvector of prices of mobile products in the j-th country and ${}^jp_\natural$ to denote the subvector of prices of immobile products. Similarly, ${}^jq_*$ will denote the vector of prices of mobile resources in the j-th country, and ${}^jq_\natural$, the vector of prices of immobile resources.

DEFINITION 2.4.1. *We say that* p^0, q^0, x^0 $(or\ {}^jp^0, {}^jq^0, {}^jx^0;\ j = 1, \cdots, J)$ *are associated whenever the saddle point inequality of Theorem* 2.3.2 *is satisfied for* p^0, q^0, *and* x^0 $(or\ {}^jp^0, {}^jq^0, {}^jx^0)$.

We say that p^0 $(or\ {}^jp^0)$ *is a price vector associated with the efficient point* x^0 $(or\ {}^jx^0)$ *whenever* p^0, q^0, x^0 $(or\ {}^jp^0, {}^jq^0, {}^jx^0)$ *are associated for some* q^0 $(or\ {}^jq^0)$.

Similarly, we say that x^0 $(or\ {}^jx^0)$ *is an efficient point associated with the price vector* p^0 $(or\ {}^jp^0)$ *whenever* p^0, q^0, x^0 $(or\ {}^jp^0, {}^jq^0, {}^jx^0)$ *are associated for some* q^0 $(or\ {}^jq^0)$.

THEOREM 2.4.1. *Let the mappings*

$$G: X \to Y, \qquad F: X \to Z$$

be concave and differentiable, and let them include transportation activities. Then, if x^0 *is world efficient, it follows that there exist price vectors,* p^0, q^0 *associated with* x^0 *such that*

$$ {}^jp^0_* = p^0_* \quad and \quad {}^jq^0_* = q^0_* \qquad for\ j = 1, 2, \cdots, J.$$

PROOF. Applying the equivalence theorem of Kuhn and Tucker (Theorem 2.3.1), the world efficiency of x^0 implies the existence of $p^0 > 0$ and q^0 such that (x^0, q^0) satisfy the saddle point inequality for the function

$$\phi(x, q) = p^0G(x) + qF(x)\ .$$

Therefore, it remains only to show that ${}^jp^0_*$ are the same for all j, and similarly that ${}^jp^0_*$ are the same for all j.

Applying Theorem 2.3.2, which states necessary and sufficient conditions for a saddle value of ϕ at (x^0, q^0), the world efficiency of x^0 implies $\phi^0_x \leqq 0$. In particular, for those components of x which refer to transportation, i.e., for those in T, it follows that,

$$\frac{\partial \phi}{\partial {}^it^k_\alpha} \equiv {}^kp^0_\alpha - {}^ip^0_\alpha \leqq 0\ , \qquad \frac{\partial \phi}{\partial {}^kt^i_\alpha} \equiv {}^ip^0_\alpha - {}^kp^0_\alpha \leqq 0\ ,$$

$$\text{for } \alpha = 1^*, \cdots, A^*;\ i, k = 1, 2, \cdots, J.$$

Similarly,

$$\frac{\partial \phi}{\partial {}^{i}t^{k}_{\beta}} = {}^{k}q^{0}_{\beta} - {}^{i}q^{0}_{\beta} \leqq 0\,, \qquad \frac{\partial \phi}{\partial {}^{k}t^{i}_{\beta}} = {}^{i}q^{0}_{\beta} - {}^{k}q^{0}_{\beta} \leqq 0\,,$$

$$\text{for } \beta = 1^{*}, \cdots, B^{*};\ i, k = 1, 2, \cdots, J.$$

It follows that

$${}^{i}p^{0}_{*} = {}^{k}p^{0}_{*}\,, \quad \text{and} \quad {}^{i}q^{0}_{*} = {}^{k}q^{0}_{*} \qquad \text{for } i, k = 1, \cdots, J\,.$$

We may call the common values p^{0}_{*} and q^{0}_{*} respectively.

Before stating the converse of this theorem, we notice that Theorem 2.4.1 asserts that world efficiency implies that the prices of mobile commodities are the same in all countries, when transportation is costless. This, of course, would be a consequence of the existence of free international trade. What is of greater interest is the converse, namely, that local efficiency with international equality of prices of mobile commodities is sufficient for world efficiency. Since free international trade is supposed to guarantee the international equality of the prices of "tradable" commodities (and free competition the local efficiency of production), this converse is very close to being a statement of the classical welfare theorem of international trade theory. This question is discussed in Section 2.5.

THEOREM 2.4.2. *Let the mappings*

$$G: X \to Y\,, \qquad F: X \to Z$$

be concave and differentiable and let them include transportation activities. Let x^{0} have subvectors ${}^{j}x^{0}$ $(j = 1, 2, \cdots, J)$, and let ${}^{j}x^{0}$ be locally efficient for each $j = 1, 2, \cdots, J$. Then, if there are ${}^{j}p^{0}$ and ${}^{j}q^{0}$ associated with ${}^{j}x^{0}$ such that

$${}^{j}p^{0}_{*} = p^{0}_{*} \quad \textit{and} \quad {}^{j}q^{0}_{*} = q^{0}_{*} \qquad \textit{for } j = 1, 2, \cdots, J\,,$$

then it follows that x^{0} is world efficient.

PROOF. For any ${}^{j}p^{0}$ and ${}^{j}q^{0}$ associated with ${}^{j}x^{0}$, we are assured by Theorem 2.3.1 that

$${}^{j}p^{0}\,{}^{j}G({}^{j}x) + {}^{j}q^{0}\,{}^{j}F({}^{j}x) \leqq {}^{j}p^{0}\,{}^{j}G({}^{j}x^{0}) + {}^{j}q^{0}\,{}^{j}F({}^{j}x^{0}) \leqq {}^{j}p^{0}\,{}^{j}G({}^{j}x^{0}) + {}^{j}q\,{}^{j}F({}^{j}x^{0})$$

$$\text{for all } {}^{j}x \geqq 0,\ {}^{j}q \geqq 0 \text{ and } j = 1, 2, \cdots, T\,.$$

From the hypotheses of the theorem we see that there is a p^{0} and a q^{0} with

$${}^{j}p^{0}_{*} = p^{0}_{*}\,, \qquad {}^{j}q^{0}_{*} = q^{0}_{*} \qquad \text{for } j = 1, 2, \cdots, J\,.$$

Substituting in the inequalities above and summing on j yields that

$$p^{0}G(x) + q^{0}F(x) \leqq p^{0}G(x^{0}) + q^{0}F(x^{0}) \leqq p^{0}G(x^{0}) + qF(x^{0})$$

$$\text{for all } x \geqq 0,\ q \geqq 0,$$

which implies the world efficiency of x^0.

2.5. *Competition and efficiency; The Classical Welfare Theorem.* The results obtained in the earlier sections characterize efficient world production in a manner free of any assumptions regarding the institutional structure of production, including transportation of commodities.

These results consist of characterizations of locally and world efficient production in terms of prices and in terms of the relationships between production in the various countries. Thus, the results constitute a solution of one of the classical welfare problems, namely, to find the set of statically efficient international specializations of production, given the technology and resource endowment of the world. A second classical welfare problem remains: Is the organization of production according to the institutions described by free trade and competition sufficient to achieve (in the static sense) efficient world production? The answer is, of course, yes; milking of the Kuhn-Tucker results to obtain a rigorous establishment of this theorem is the task of this section.

Production, including transportation, is carried out by firms, each firm comprising one basic activity; taking prices as given to it independent of its actions, each firm strives to maximize profits. Prices are determined by (unspecified) market interactions, so that the familar conditions for competitive equilibrium are fulfilled, to wit: Each firm is maximizing its profits, going out of production if the maximum is negative. In equilibrium no firm earns positive profits, and the value in equilibrium prices of unused resources is zero. We have, in addition, the conditions implied by the institutional fact that prices are nonnegative and the technological facts that activity variables are nonnegative and that the rate of resource utilization cannot exceed the limited availability of resources.[13] These conditions fulfill precisely the conditions of Theorem 2.3.2. These conditions are necessary and sufficient for a saddle value of the Lagrangean function $\phi(x, q)$ and hence for the world efficiency of production.

The sufficiency of the competitive organization of production for (static) efficiency is summed up in the following statement of the classical welfare theorem of international trade theory.

THEOREM 2.5.1. *Let the mappings*

$$G: X \to Y, \qquad H: X \to Z$$

be concave and differentiable. Let the world technology be organized

[13] See P. A. Samuelson, *Foundations of Economic Analysis* (Cambridge, Massachusetts: Harvard University Press, 1948), pp. 73–89.

according to the rules of perfect competition. Then, if x is an activity vector which is carried out in competitive equilibrium, it follows that x is also world efficient.

3. FACTOR PRICES

3.1. *Introduction.* We are now ready to turn to the factor price equalization problem. As we have just seen, world efficiency of production implies (and hence for any equilibrium point of a freely trading competitive international economy it follows) that the prices of mobile commodities can be taken to be the same in all countries. Because of the technological interconnections among all commodities in any one country, if not by reason of earlier theorems, it is plausible to expect that the prices of immobile commodities be related to the prices of mobile ones. Specifically, the question is whether world efficiency implies that the prices of internationally immobile commodities must also be equal in different countries.

Samuelson[14] has shown that if:

(a) all products are mobile and all resources immobile;
(b) all countries have the same technology;
(c) constant returns to scale prevail in all lines of production;
(d) there is no joint production;
(e) a certain regularity condition, Lemma 3.3.1 below, (generally interpreted in terms of non-reversibility of factor intensities in production) is satisfied by the production functions;
(f) the number of products is at least as large as the number of resources;
(g) free international trade and perfect competition prevail;

then the (real) prices of the factors of production will be the same in all countries.

We shall here weaken assumptions (a), (b), (c), and (e), and under the weaker conditions derive necessary as well as sufficient conditions for factor price equalization.

Before modifying the model of Part 2 to reflect the assumption that there is no joint production we give, for the record, one obvious necessary condition for factor price equalization.

It is clear that the equalization of factor prices at a world efficient point x implies that all countries have the same effective resource

[14] P. A. Samuelson, "International Trade and the Equalization of Factor Prices," *Economic Journal*, Vol. LVIII (June, 1948), pp. 163-184; "International Factor Price Equalization Once Again," *Economic Journal*, Vol. LIX (June, 1949), pp. 181-197; "Prices of Factors and Goods in General Equilibrium," *The Review of Economic Studies*, Vol. XXI (1), No. 54 (1953-1954), pp. 1-20.

limits. This is readily seen from the fact that among the necessary conditions for a saddle-value of the Lagrangean function of Theorem 2.3.2 are that

$$\phi^0_{j_0}{}^j q \equiv {}^j q({}^j c - {}^j H({}^j x)) = 0 \qquad {}^j q \geqq 0 ,$$

and

$$\phi^0_{j_0} \equiv ({}^j c - {}^j H({}^j x)) \geqq 0 .$$

Now suppose

$${}^j q = {}^{j'} q \qquad \text{for all } j, j' = 1, 2, \cdots, J ,$$

and suppose

$${}^j q_i > 0 \qquad \text{for some } j \text{ and some coordinate } i.$$

This implies ${}^j q_i > 0$ for all j and hence that

$${}^j c_i = {}^j H_i({}^j x) \qquad \text{for all } j.$$

3.2. *Model of technology assuming no joint production.* The absence of joint production permits the identification of each activity with the single product which it produces. We can then distinguish without ambiguity the inputs or resources used in that activity. The following notations take advantage of this fact.

Recall the space of resources,

$$Z = {}^{(1)}Z \otimes \cdots \otimes {}^{(J)}Z ,$$

and of products,

$$Y = {}^{(1)}Y \otimes \cdots \otimes {}^{(J)}Y .$$

Denote by ${}^{(j)}z^\delta_\alpha$ the quantity (flow) of the α-th resource which is used in the production of the δ-th product in the j-th country.

Let ${}^jZ^D$, $j = 1, 2, \cdots, J$, denote the Euclidean space whose elements are the vectors with components ${}^jz^\delta_\alpha$, where δ runs through all products (the coordinates of ${}^jY$) and α runs through all resources (the coordinates of ${}^jZ$). Let

$$Z^D = {}^{(1)}Z^D \otimes \cdots \otimes {}^{(J)}Z^D .$$

Production can be represented by mappings

$$\hat{G}: Z^D \to Y , \qquad \hat{H}: Z^D \to Z .$$

The component ${}^jg^\delta$ of $\hat{G}$ depends only on those coordinates ${}^jz^\delta_\alpha$ of Z^D for which $i = j$ and $\varepsilon = \delta$. The mapping $\hat{H}$ has the form

$$\hat{H}(z) = \sum_i z^i , \qquad z \in Z^D .$$

For a given vector $c \in Z$, write

$$\hat{F} = c - \hat{H} .$$

This formulation is, of course, the traditional production function representation of technology, except for the formality that the functions are defined with all inputs as arguments in all functions. However, we write as different from zero only those components on which the values of the functions actually depend. Then, the arguments of the δ-th production function in the j-th country are the inputs used in producing commodity δ in country j.

Moreover, this model is a special case of the Kuhn-Tucker non-linear activity analysis model discussed in Part 2. Identify the space Z^D with the space X, the mapping $\hat{G}$ with the mapping G, and the mapping $\hat{H}$ with the mapping H. The spaces Y and Z are the spaces Y and Z of Part 2. With these identifications, the following theorem is true.

THEOREM 3.2.1. *Let the mapping $\hat{G}: Z^D \to Y$ be concave [$\hat{H}$ is concave, since it is linear] and differentiable. Then, the world efficiency of the point $z \in Z^D$ is equivalent to the existence of a vector $p > 0$ such that z and some vector $q \geqq 0$ satisfy the saddle point inequality for the function*

$$\phi(z, q) = p\tilde{G}(z) + q\tilde{F}(z) .$$

We now define precisely what we mean by factor price equalization.

DEFINITION 3.2.1. *Let $z \in Z^D$ be a world efficient point, and let p and q be price vectors associated with it. We say that factor prices are equalized whenever*

$$^{j}p_{i} = {}^{i}p_{i} \quad \text{and} \quad {}^{j}q_{i} = {}^{i}q_{i} \qquad \text{for all } i, j = 1, 2, \cdots, J ,$$

for arbitrary choice of price vectors p and q associated with z.

The problem at hand, then, is to find such necessary and sufficient conditions that the price vectors associated with a world efficient point satisfy the condition for factor price equalization.

3.3. *Two countries, two mobile products, two immobile resources.* In order to avoid the notational complexities involved in the many-dimensional general case, we treat first the simplest non-trivial case, that involving just two mobile products, two immobile resources, and two countries. Thus,

ASSUMPTION 3.3.1. (1) $A^1 = B^* = 0$. (2) $A^* = B^1 = 2$, (3) $J=2$.

This assumption permits the following notational simplification.

DEFINITION 3.3.1. *Let*

$$x_i = {}^1z_{i1}^{1*}, \qquad u_i = {}^2z_{i1}^{1*}, \qquad (i = 1, 2)$$
$$y_i = {}^1z_{i1}^{2*}, \qquad v_i = {}^2z_{i1}^{2*}, \qquad (i = 1, 2)$$
$$c_i = {}^1c_{i1}, \qquad d_i = {}^2c_{i1}, \qquad (i = 1, 2)$$
$$f(\ ,\ ) = {}^1g^{1*}(\ ), \quad h(\ ,\ ) = {}^2g^{1*}(\ ),$$
$$g(\ ,\ ) = {}^1g^{2*}(\ ), \quad k(\ ,\ ) = {}^2g^{2*}(\ ).$$

Here x, y, f, and g refer to country 1, while u, v, h, and k refer to country 2. The subscript refers to the input commodity, the letter to the output comodity. Thus, x_1 is the amount of the first input used in producing commodity 1 in country 1; u_1 denotes the corresponding quantity in country 2. Further, note

$$f_1^0 = \frac{\partial f(x_1^0, x_2^0)}{\partial x_1}, \qquad f_{12}^0 = \frac{\partial^2 f(x_1^0, x_2^0)}{\partial x_1 \partial x_2},$$

and similarly for the derivatives of other functions.

Using the new notation, the Lagrangean function ϕ of Theorem 3.2.1 becomes (writing x for the vector of inputs),

$$\phi(x, \bar{q}, \bar{r}) = \begin{pmatrix} p_1 f(x_1, x_2) + p_2 g(y_1, y_2) + \bar{q}_1[c_1 - (x_1 + y_1)] + \bar{q}_2[c_2 - (x_2 + y_2)] \\ p_1 h(u_1, u_2) + p_2 k(v_1, v_2) + \bar{r}_1[d_1 - (u_1 + v_1)] + \bar{r}_2[d_2 - (u_2 + v_2)] \end{pmatrix}.$$

We are, as has been pointed out above, interested in the prices associated with a world efficient point. Theorem 3.2.1 applied to the case at hand informs us that world efficiency is equivalent to the attainment of a saddle-value of the function $\phi(x, \bar{q}, \bar{r})$. Further, we can apply Theorem 2.3.2 to obtain some conditions implied by such a saddle value. Thus,

THEOREM 3.3.1. *If the function $\phi(\ ,\ ,\ )$ has a saddle-value at the point $(x, \bar{q}, \bar{r})$, then it follows that*

$$p_1 f_i - \bar{q}_i \leqq 0, \qquad p_1 h_i - \bar{r}_i \leqq 0,$$
$$p_2 g_i - \bar{q}_i \leqq 0, \qquad p_2 k_i - \bar{r}_i \leqq 0,$$
$$x_i + y_i \leqq c_i, \qquad u_i + v_i \leqq d_i,$$

where $i = 1, 2$. Since $p_1 > 0$ and $p_2 > 0$, we can perform the following normalization without changing the inequalities in any essential way. Let

$$q_i = \frac{\bar{q}_i}{p_1}, \qquad p = \frac{p_2}{p_1}, \qquad r_i = \frac{\bar{r}_i}{p_1}, \qquad (i = 1, 2).$$

We are here interested only in the case of an interior saddle-value.[15] Thus, we assume that all relevant quantities are positive.

DEFINITION 3.3.2. *The world efficient point x is called unspecialized if the components $x_1, x_2, y_1, y_2, u_1, u_2, v_1, v_2$, and q_1, q_2, r_1, and r_2 are all positive.*

ASSUMPTION 3.3.2. *The point x is unspecialized world efficient.*

In the light of Theorem 2.3.2, this condition implies that the inequalities of Theorem 3.3.1 are in fact equalities. Thus, the unspecialized world efficiency of x implies that there exist solutions p, $q_1, q_2, x_1, x_2, y_1, y_2$, and $p, r_1, r_2, u_1, u_2, v_1, v_2$ respectively to the equation systems

$$\begin{aligned} pf_i - q_i &= 0\,, & ph_i - r_i &= 0\,, \\ g_i - q_i &= 0\,, & k_i - r_i &= 0\,, \qquad (i = 1, 2) \\ x_i + y_i &= c_i\,, & u_i + v_i &= d_i\,. \end{aligned} \tag{3.3.1}$$

Now the first column of equations (those referring to country 1) is a system which determines x_1, x_2, y_1, y_2, q_1, and q_2 in terms of p, c_1, and c_2; the second column determines u_1, u_2, v_1, v_2, r_1, and r_2 in terms of p, d_1, and d_2, provided an unspecialized world efficient point exists for the specified p, c_1, c_2, d_1, and d_2. The existence of such a point implies the existence of solutions of these equations. The factor price equalization problem reduces to the question, "Do these solutions have the property that

$$q_1 = r_1\,, \quad \text{and} \quad q_2 = r_2$$

for an arbitrary choice of q and r associated with the unspecialized world efficient point?"

Thus, the first problem to settle is that of the existence of world efficient points. This can be guaranteed by one of a variety of conditions. The simplest course here is to assume the existence of an unspecialized, world efficient point.

ASSUMPTION 3.3.3. *There are unspecialized world efficient points.*

Having assumed the existence of unspecialized world efficient points, and hence of the solutions to the equations at hand, we are prepared to discuss the nature of these solutions.

It is clear that if we are to have factor price equalization two further conditions must hold:

[15] The effect of this limitation on the generality of the analysis in many dimensions is discussed in Section 3.7.

(1) The solution for q_1, q_2 and r_1, r_2 cannot depend on c_1, c_2 and d_1, d_2 respectively.

(2) The solutions must be unique.

It is evident that if the solutions depend on the resource endowments c_1, c_2, and d_1, d_2, then we cannot have factor price equalization for all $p > 0$ and all c_1, c_2, d_1, d_2.

Second, if at a world efficient point there exist two distinct solutions for r_1, r_2 say, then there is at least one choice of price vectors which has the property that

$$q_1 \neq r_1 , \quad \text{and} \quad q_2 \neq r_2 .$$

Consequently, we must look for conditions under which there exist unique solutions q_1, q_2 and r_1, r_2 of the equations (3.3.1) which do not depend on c_1, c_2 and d_1, d_2. Thus, we may drop the equations involving c_1, c_2, d_1, d_2 and investigate the remaining system, namely the system

$$\text{(3.3.2)} \qquad \begin{aligned} pf_i - q_i &= 0 , & ph_i - r_i &= 0 , \\ g_i - q_i &= 0 , & k_i - r_i &= 0 , \end{aligned} \qquad (i = 1, 2) .$$

Now, the system of equations (3.3.2) falls into two sub-systems, those referring to country 1, and those referring to country 2. Each subsystem involves seven variables; for country 1 they are $p, x_1, x_2, y_1, y_2, q_1, q_2$; for country 2, $p, u_1, u_2, v_1, v_2, r_1, r_2$. The functions f, g and h, k are in general all different. The question is, "What conditions on these functions will imply solutions q_1, q_2 and r_1, r_2 such that for each $p > 0$

$$q_1 = r_1 , \quad \text{and} \quad q_2 = r_2 \ ?"$$

That is, under the Assumptions 3.3.1, 3.3.2, and 3.3.3, we wish to find conditions which are sufficient for factor price equalization. The following theorem states these conditions. The interpretation of them will be discussed below.

THEOREM 3.3.2. *Let $V(\ ,\ )$ and $W(\ ,\ )$ be continuously differentiable functions of two variables such that*[16]

$$\text{(1)} \qquad V_1 \not\equiv 0 ,$$

$$\text{(2)} \qquad \begin{vmatrix} V_1 & V_2 \\ W_1 & W_2 \end{vmatrix} \not\equiv 0 .$$

If for such V, W the functions f, g, h, and k satisfy

[16] The notation $\not\equiv$ is read "identically different from." As above, the subscript i denotes differentiation with respect to the i-th variable. The symbol $|A|$ denotes the determinant of A.

$$V(f_1, f_2) \equiv 0\,, \qquad V(h_1, h_2) \equiv 0\,,$$
$$W(g_1, g_2) \equiv 0\,. \qquad W(k_1, k_2) \equiv 0\,,$$

identically in their arguments, then equations (3.3.2) *have for each* $p > 0$ [17] *a unique solution* q_1, q_2, r_1, r_2 *with the property that*

$$q_1 = r_1\,, \quad \text{and} \quad q_2 = r_2\,,$$

i.e., the conditions stated are sufficient for factor price equalization.

PROOF. Rewrite equations (3.3.2) to eliminate q_1, q_2, r_1, and r_2. Thus,

$$pf_i - g_i = 0\,, \qquad ph_i - k_i = 0\,. \qquad (i = 1, 2)\,.$$

From the hypotheses of the theorem we have in addition the equations

$$V(f_1, f_2) \equiv 0\,, \qquad V(h_1, h_2) \equiv 0\,,$$
$$W(g_1, g_2) \equiv 0\,, \qquad W(k_1, k_2) \equiv 0\,.$$

Consider the sub-system of four equations referring to country 1. It follows from Assumption 3.3.3 that this sub-system has for each $p > 0$ a solution $f_1^0, f_2^0, g_1^0, g_2^0$. But, considering the derivatives f_1, f_2, g_1, g_2 as the variables, the equations referring to country 2 are the same as those referring to country 1. Consequently, any solution for the equations of country 1 must also be a solution of those referring to country 2. That is, if $f_1^0, f_2^0, g_1^0, g_2^0$ constitute a solution for country 1, then $h_1^0, h_2^0, k_1^0, k_2^0$ where

$$h_1^0 = f_1^0\,, \qquad h_2^0 = f_2^0\,, \qquad k_1^0 = g_1^0\,, \qquad k_2^0 = g_2^0$$

constitute a solution for country 2.

It remains to show that there can be just one solution for country 1. Consider the system

$$pf_i - g_i = 0\,, \qquad V(f_1, f_2) = 0\,, \qquad W(g_1, g_2) = 0\,, \qquad (i = 1, 2)\,.$$

Its Jacobian is

$$\begin{vmatrix} p & 0 & -1 & 0 \\ 0 & p & 0 & -1 \\ V_1 & V_2 & 0 & 0 \\ 0 & 0 & W_1 & W_2 \end{vmatrix}.$$

Interchanging the first and third column yields the determinant

[17] By considering p as a given quantity which can have any positive value, we provide, as remarked above, for the totality of possible demand configurations.

$$\begin{vmatrix} -1 & 0 & p & 0 \\ 0 & p & 0 & -1 \\ 0 & V_2 & V_1 & 0 \\ W_1 & 0 & 0 & W_2 \end{vmatrix},$$

which has the property that its successive principal minors

$$D_1 = -1, \qquad D_2 = \begin{vmatrix} -1 & 0 \\ 0 & p \end{vmatrix},$$

$$D_3 = \begin{vmatrix} -1 & 0 & p \\ 0 & p & 0 \\ 0 & V_2 & V_1 \end{vmatrix}, \quad D_4 = \begin{vmatrix} -1 & 0 & p & 0 \\ 0 & p & 0 & -1 \\ 0 & V_2 & V_1 & 0 \\ W_1 & 0 & 0 & W_2 \end{vmatrix}$$

are identically non-vanishing in consequence of the hypotheses concerning V and W. Thus, the hypotheses of the following lemma are satisfied, and hence its conclusions apply and the uniqueness of the solution $f_1^0, f_2^0, g_1^0, g_2^0$ is proved. These hypotheses are assumed by Samuelson and interpreted in the two product, two factor case to require non-reversal of relative factor intensities.

LEMMA 3.3.1. *Let $f^i(w_1, w_2, \cdots, w_n)h_2 \leqq w_2 \leqq c_i$ $(i = 1, 2, \cdots, n)$ be n functions of n real variables and let the f^i be continuously differentiable. A sufficient condition that the equations*

$$f^i(w_1, \cdots, w_n) = a^i$$

have at most one solution $w_1^0, w_2^0, \cdots, w_n^0$ is that there exist a numbering of the functions and variables such that the successive principal minors

$$D_1 = \frac{\partial f^1}{\partial w_1}, \; D_2 = \begin{vmatrix} \frac{\partial f^1}{\partial w_1} & \frac{\partial f^1}{\partial w_2} \\ \frac{\partial f^2}{\partial w_1} & \frac{\partial f^2}{\partial w_2} \end{vmatrix}, \cdots, D_n = \begin{vmatrix} \frac{\partial f^1}{\partial w_1} & \cdots & \frac{\partial f^1}{\partial w_n} \\ \vdots & & \vdots \\ \frac{\partial f^n}{\partial w_1} & \cdots & \frac{\partial f^n}{\partial w_n} \end{vmatrix}$$

be identically non-vanishing in $w_1, \cdots, w_n$.[18]

Returning to the proof of Theorem 3.3.2, we now have a unique solution $f_1^0, f_2^0, g_1^0, g_2^0$ of the equations for country 1, and similarly for country 2 with $h_1^0 = f_1^0$, $h_2^0 = f_2^0$, $k_1^0 = g_1^0$, $k_2^0 = g_2^0$. Substituting these values in equations (3.3.2) we find solutions for q_1, q_2, r_1, r_2 which must then have the property that

[18] A proof of this lemma is to be found in P. A. Samuelson, "Prices of Factors and Goods in General Equilibrium," *The Review of Economic Studies*, No. 54 (1953-54), p. 16.

$$q_1 = r_1\,, \qquad q_2 = r_2 \qquad \text{for all } p > 0\,.$$

Before discussing the meaning of the conditions of Theorem 3.3.2, we present a necessary condition for factor price equalization. This condition strongly suggests that the sufficient condition just stated is the weakest possible sufficient condition (apart from trivial modifications).

Consider equations (3.3.1). Under Assumptions 3.3.1 to 3.3.3 these equations have for each $p > 0$ and every c_1, c_2, d_1, d_2 a positive solution $x_1^0, x_2^0, y_1^0, y_2^0, u_1^0, u_2^0, v_1^0, v_2^0, q_1^0, q_2^0, r_1^0, r_2^0$.

Consider the sets of positive points $x = (x_1, x_2, y_1, y_2, u_1, u_2, v_1, v_2)$ which are for some $p > 0$, c_1, c_2, d_1, d_2 involved in a solution of equations (3.3.1). These points are world efficient, forming a subset of the world efficient points characterized by the fact that both countries produce both commodities.

Further, for every point x having the properties that

(1) for some $p > 0$, c_1, c_2, d_1, d_2, x is world efficient, and

(2) x is strictly positive,

it follows that x belongs to the set in question. This set is the set of unspecialized world efficient points.

In general the set of unspecialized world efficient points is a "curve" (possibly a domain) in the activity space. It is necessary for factor price equalization that the production function $f, g, h,$ and k, satisfy relations of functional dependence among its derivatives on the set of unspecialized world efficient points.

THEOREM 3.3.3. *For every solution q_1, q_2, r_1, r_2 which equations* (3.3.1) *have for each $p > 0$ and c_1, c_2, d_1, d_2, let it be true that*

$$q_1 = r_1\,, \quad \text{and} \quad q_2 = r_2\,.$$

Then, it follows that there exist functions V and W of two variables, such that $f, g, h,$ and k satisfy

$$V(f_1, f_2) = 0\,, \qquad V(h_1, h_2) = 0\,,$$
$$W(g_1, g_2) = 0\,, \qquad W(k_1, k_2) = 0\,,$$

for x, an unspecialized world efficient point.

PROOF. That

$$q_1 = r_1\,, \quad \text{and} \quad q_2 = r_2$$

for all $p > 0$ and all c_1, c_2, d_1, d_2 implies

(1) that the solutions $q_1, q_2, r_1,$ and r_2 are independent of c_1, c_2, d_1, d_2, and

(2) that q_1 and q_2 are uniquely determined by p, given f and g. Thus, there are functions $\alpha(\ )$ and $\beta(\ )$ such that

$$q_1 = \alpha(p), \qquad q_2 = \beta(p), \qquad p > 0$$

and, from the hypotheses, r_1 and r_2 are the same functions of p

$$r_1 = \alpha(p), \qquad r_2 = \beta(p), \qquad p > 0.$$

Substitution in equations (3.3.1) yields that

$$f_1 = \alpha(p)/p = h_1, \qquad g_1 = \alpha(p) = k_1,$$
$$f_2 = \beta(p)/p = h_2, \qquad g_2 = \beta(p) = k_2,$$

which implies the existence of the functions V and W asserted in the conclusion of the theorem.

3.4. *Analysis of the conditions for factor price equalization.* The sufficient condition of Theorem 3.3.2 was expressed by requiring the existence of functions V and W of two variables such that the derivatives of the production function, i.e., the marginal productivities, satisfied the conditions

$$V(f_1, f_2) \equiv 0, \qquad V(h_1, h_2) \equiv 0,$$
$$W(g_1, g_2) \equiv 0, \qquad W(k_1, k_2) \equiv 0,$$

identically in the respective inputs. These conditions obviously constitute restrictions imposed on the production functions. We now explore the nature of these restrictions.

The question has two parts. First, what restriction is imposed on a production function $f(x_1, x_2)$ by the requirement that it satisfy $V(f_1, f_2) \equiv 0$? Second, what additional restriction is imposed by requiring that both f and h satisfy the same condition?

Consider the first question, which involves the concept of functional dependence. The requirement that the derivatives of the production function satisfy the condition

$$V(f_1, f_2) \equiv 0$$

for a non-trivial differentiable function V amounts to requiring that the functions $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$, i.e., the marginal productivity functions, be dependent. More precisely,

DEFINITION 3.4.1. *Let $u(x, y)$, $v(x, y)$ each be functions of two variables having continuous first partial derivatives. We say u and v are functionally dependent if, for each point (x, y) of some domain D of the (x, y) plane, u and v satisfy*

$$V(u, v) = 0,$$

where V is a differentiable function of two variables such that V_u and V_v are never simultaneously zero at any point (u, v).

The well-known necessary and sufficient condition that two (regular) functions be functionally dependent is that their Jacobian should vanish identically in some domain.[19]

Applying this theorem on functional dependence to the partial derivatives of the production function, we find that,

THEOREM 3.4.1. *The function* $f(x_1, x_2)$ *satisfies* $V(f_1, f_2) \equiv 0$ *for some* V *if, and only if,* f *satisfies the differential equation*

$$f_{11}f_{22} - f_{12}^2 = 0 .$$

This differential equation is well known; it characterizes the class of developable surfaces.[20]

A surface $S(x, y, z) = 0$ in the (x, y, z) space is called developable if it can be rolled out (or developed) on a plane without stretching or contracting any part of it.

For example, a cylinder is developable; a sphere is not. Developable surfaces are classified into three types, namely, (a) cones, (b) cylinders, and (c) tangent surfaces of skew or twisted curves. (A curve in space is called *skew* or *twisted* if it cannot be contained in a two dimensional plane.)

It is true of a developable surface that through any point there is a line segment which is contained in the surface. We may use this fact to look further into the relationship between the function $V(\ ,\ )$ and the surface. Following a suggestion of Koopmans, we write the line through the point (x^0, x^0, y^0), where $z^0 = u(x^0, y^0)$ as

$$z(t) \equiv z^0 + \beta(x^0, y^0)t , \quad x(t) \equiv x^0 + \alpha(x^0, y^0)t , \quad y(t) \equiv y^0 + \gamma(x^0, y^0)t .$$

The condition

$$\frac{d^2}{dt^2}\{z(t) - u(x(t), y(t))\}_{t=0} = 0$$

gives us that

$$(\alpha^0)^2u_{11}^0 + \alpha^0\gamma^0(u_{12}^0 + u_{21}^0) + (\gamma^0)^2u_{22}^0 = 0 ,$$

where the superscript 0 represents the arguments (x^0, y^0) of α, γ, and the u_{ij}. Rewriting the above we have,

$$[\alpha^0, \gamma^0]\begin{bmatrix} u_{11}^0 & u_{12}^0 \\ u_{21}^0 & u_{22}^0 \end{bmatrix}\begin{bmatrix} \alpha^0 \\ \gamma^0 \end{bmatrix} = 0 .$$

Since $u_{12} = u_{21}$ and the rank of

[19] R. S. Burington and C. C. Torrance, *Higher Mathematics* (New York: McGraw-Hill, 1939), pp. 140–141.

[20] L. P. Eisenhart, *An Introduction to Differential Geometry* (Princeton: Princeton University Press, 1947), pp. 53 ff

$$\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}$$

is 1, the equation above reduces to

$$[\alpha^0, \gamma^0]\begin{bmatrix} u_{11}^0 & u_{12}^0 \\ u_{21}^0 & u_{22}^0 \end{bmatrix} = 0 ,$$

where the rank of $[u_{ij}^0]$ is assumed different from zero.

Differentiating

$$V(u_1, u_2) = 0$$

with respect to x and y yields that

$$[V_1, V_2]\begin{bmatrix} u_{11}^0 & u_{12}^0 \\ u_{21}^0 & u_{22}^0 \end{bmatrix} = 0 .$$

Now, the direction of $[\alpha^0, \gamma^0]$ is the direction of the projection on the (x, y) plane of the line through (x^0, x^0, y^0). The argument above shows that this direction is given by the values of V_1, V_2. These values are those that the functions $V_1(\ , \ )$ and $V_2(\ , \ )$ take when the marginal productivities u_1^0 and u_2^0 at the point (x^0, y^0) are inserted as arguments.

Consider now the second question: "What additional restriction is imposed by requiring that corresponding (to the same product) production functions in the different countries satisfy the same condition $V(\ , \ ) = 0$ on their partial derivatives?" To translate the condition into words, the condition that $u(x, y)$ and $v(r, s)$ are such that $V(u_1, u_2)=0$ and $V(v_1, v_2) = 0$ means that given the value of one derivative of u, the other is determined, and that the same functional dependence exists between the derivatives of v.

It seems that there is little more one can conclude from these conditions alone. In order to obtain more specific answers we shall impose the following additional conditions on the functions $u(\ , \ )$ and $v(\ , \ )$.

ASSUMPTION 3.4.1. (1) *The production surface $z = u(x, y)$ is given by a single-valued function $u(x, y)$ defined for all $x \geqq 0$, $y \geqq 0$, and continuously differentiable everywhere except possibly at the origin.*

(2) *For all $x \geqq 0$, $u(x, 0) = 0$, and for all $y \geqq 0$, $u(0, y) = 0$.*

We ask first what these additional conditions enable us to conclude about the developable surface $z = u(x, y)$. As the following theorem states, these conditions imply that the surface is part of a cone and, hence, that the production function is homogeneous of first degree, i.e., affords constant returns to scale.

THEOREM 3.4.2. *Let $u(x, y)$ satisfy Assumption* 3.4.1; *then,* $u_{11}u_{22} - u_{12}^2 = 0$ *implies that the surface given by* $z = u(x, y)$ *is a cone.*

PROOF: As we have seen, the condition

$$u_{11}u_{22} - u_{12}^2 = 0$$

characterizes the developable surfaces. As we remarked above, the developable surfaces are completely classified; such a surface is either a cone, a cylinder, or the tangent surface to a space curve. We consider the cases separately.

(1) If the surface is a cone, the theorem is true.

(2) If the surface is a cylinder, then it coincides with the (x, y) plane, for our conditions state that the cylinder intersects the (x, y) plane only in the x axis. Let x^0 be a point of the x axis. Then there is a non-vertical line l_0 through x^0, contained in the cylinder. Take $x' \neq x^0$ on the x axis. Then there is a line l_1, parallel to l_0 through x' and contained in the cylinder. Thus, there is a non-vertical plane containing the x axis and contained in the cylinder. The case is similar for the y axis. Since these planes are non-vertical, either they coincide in the (x, y) plane or they intersect. But the differentiability conditions, together with the requirement that $u(x, y)$ be single-valued and everywhere defined, clearly rule out the case that the planes intersect. Hence the surface is the (x, y) plane and a degenerate cone.

This situation is geometrically obvious and hence only a sketch of the proof is given.

(3) Let $z = u(x, y)$ give the tangent surface of a space curve. Let $r(s) = a(s)i + b(s)j + c(s)k$ be a space curve, where i, j, and k are unit normal vectors of our space, and $a(s)$, $b(s)$, and $c(s)$ are continuously differentiable functions of arc length s.

We may write the surface

$$r^* = r(s) + \lambda t(s)$$

for s in some interval $s_0 \leqq s \leqq s_1$, $-\infty < \lambda < \infty$, and where

$$t(s) = a'(s)i + b'(s)j + c'(s)k ,$$

where $'$ denotes differentiation with respect to the parameter s. Thus,

$$r^* = [a(s) + \lambda a'(s)]i + [b(s) + \lambda b'(s)]j + [c(s) + \lambda c'(s)]k$$

for some λ and s. Hence:

$$b(s) + \lambda b'(s) = 0 , \qquad c(s) + \lambda c'(s) = 0 .$$

Consider an interval of the x axis containing the point xi. Suppose

that this interval is the image of a curve in the (s, λ) plane, so that λ is a single-valued function of s for $s' \leqq s \leqq s''$.[21]

We have then the differential equations

$$b(s) + \lambda(s)b'(s) = 0\,, \qquad c(s) + \lambda(s)c'(s) = 0\,,$$

which imply

$$b(s) = mc(s)\,, \qquad m \text{ being a constant.}$$

Therefore,

$$r(s) = a(s)i + c(s)[mj + k] \qquad \text{for } s' \leqq s \leqq s''\,,$$

which tells us that the curve $r(s)$ lies in a plane through the x axis.

Similarly, we arrive at the result that the curve lies in a plane through the y axis. Again, as in the case of the cylinder, it is clear that these planes must coincide in the (x, y) plane. (Otherwise the surface would not be defined everywhere, or would not satisfy the regularity conditions imposed.)

We return now to the second question, "What additional restriction is imposed by requiring that corresponding (to the same product) production functions in different countries should satisfy the same condition $V(\ ,\ ) = 0$ on their partial derivatives?"

Let $u(x, y)$ and $v(x, y)$ be homogeneous of first degree. We then write

$$u(x, y) = xu(1 \cdot y/x) = x\bar{u}(t), \qquad \text{where } t = y/x.$$

Hence

$$u_1 = -t\bar{u}'(t) + \bar{u}(t)\,, \qquad u_2 = \bar{u}'(t)\,,$$

and

$$u_1 = -tu_2 + \bar{u}(t)\,.$$

Similarly, if $v(x, y) = x\bar{v}(t)$, then $v_1 = -tv_2 + \bar{v}(t)$. Now, if v_1 is to be the same function of v_2 as u_1 is of u_2 we must, for each t, have $\bar{v}(t) = \bar{u}(t)$, which implies $v(x, y) = u(x, y)$. Thus we are able to conclude the following:

THEOREM 3.4.3. *Let the functions f, g, h, and k satisfy the conditions of Theorem* 3.3.2, *and in addition let them each satisfy Assumption* 3.4.1. *Then*

$$f = h\,, \qquad g = k\,,$$

[21] If this is not true, then the curve must have a vertical tangent at every point in the interior, for our conditions require that the surface intersect the (x, z) plane only in the x axis.

and these functions are homogeneous of first degree.

We are now able to see the relationship of our conditions to those of Samuelson.

The proof of Theorem 3.3.3 shows us that the existence of functional dependences V, W, among the marginal productivity functions ensures the satisfaction of Samuelson's regularity condition (hypotheses of Lemma 3.3.1).

If, in addition, we make Assumption 3.4.1, the main point of which is that some of each input is required to produce a positive output of any commodity, then Samuelson's conditions (b) and (c) above are also implied, and we obtain Samuelson's model as a specialization of the one given here.

More precisely, the relationships among the various conditions can be made clear as follows. We first list the following statements:

A. Production functions satisfy conditions of functional dependence among marginal productivity functions.

B. The same functional dependences are satisfied in both countries.

C. Production functions afford constant returns to scale.

D. Production functions are identical in both countries.

E. A positive amount of every input is needed to produce any output.

F. Factor price equalization prevails.

Then, Theorems 3.3.2 and 3.3.3 together assert that

$$F \Longleftrightarrow (A \text{ and } B).$$

Theorem 3.4.2 asserts

$$(A \text{ and } E) \text{ implies } C.$$

Theorem 3.4.3 asserts

$$(C \text{ and } B) \text{ implies } D,$$

from which it is clear that

$$(F \text{ and } E) \text{ implies } (C \text{ and } D).$$

Moreover, we see that

$$(C \text{ and } D) \text{ implies } F,$$

since C implies A, and (A and D) implies B.

This is, of course, the essential content of Samuelson's factor price equalization theorem.

3.5. *A generalization to the case of by-products.* We shall indicate how the assumption of no joint production can be relaxed slightly.

The case of by-products can be handled with no extra work. The key fact is that when commodities are produced as by-products it still remains true that inputs can be identified with the particular processes producing sets of by-products.

Let the index δ run through the set of processes (a process produces either a single product or a set of by-products). Let the by-products of the δ-th process be indexed $\delta_1, \delta_2, \cdots, \delta_k$. As in Section 3.2, let i,δ denote the quantity of the δ-th input in the δ-th process. Again considering the case of two processes, let $w_{i,k}$ denote the i-th by-product of the k-th process, and write

$$w_{i1} = f^i(x_1, x_2)\,, \quad w_{i2} = g^i(y_1, y_2)\,, \quad x_i + y_i = c_i\,, \quad (i = 1, 2)\,,$$

where x_1 and x_2 denote the quantities of inputs used in process 1, and y_1, y_2 denote the quantities of inputs used in process 2. We have, of course, written only the equations for country 1. It follows then that the equations obtained in place of (3.3.2) are (for country 1),

$$p_1 f_i^1 + p_2 f_i^2 - q_i = 0\,, \qquad p_3 g_i^1 + p_4 g_i^2 - q_i = 0\,, \qquad (i = 1, 2)\,,$$

and similarly for country 2. Here p_1, p_2, p_3, p_4 are the prices of the outputs $w_{11}, w_{12}, w_{21}, w_{22}$ respectively. It is evident that the same line of argument used in the case already treated applies here.[22]

3.6. *Generalization to the case of many countries, many products, many resources.* It remains to generalize the results obtained in the "2 × 2 × 2" case (two countries, two mobile products, two immobile resources) to the case of many countries, many products, and many resources, when both internationally mobile and internationally immobile ones may be present. The line of attack used in the "2×2×2" case is also used in the higher dimensional case; therefore, we do not give the remaining analysis in complete detail, but rather sketch it. We carry out the generalization in steps. The first step is to consider a larger number of mobile products, of immobile resources, and of countries. The second step will be to consider the case in which there are only immobile products and mobile resources. The third step is to allow both mobile and immobile products together with immobile resources, and finally, the last step is to permit mobile and immobile products together with mobile and immobile resources.

After completing this generalization we will discuss the interpretation of the conditions for factor price equalization in many dimensions.

[22] Certain other extensions of the analysis in the direction of allowing $F(z)$ to be a linear function rather than merely a sum are clearly possible. However, in that case the vector q would no longer have the same interpretation.

We shall then discuss the relevance of Assumption 3.3.2, which, in effect, assumes that every country produces every commodity. This assumption can be weakened considerably, as we shall see.

Returning now to the model and notations of Section 3.2, recall that Theorem 3.2.1 characterized world efficiency of production in terms of saddle-values of the function $\phi(z, q) = p\hat{G}(z) + q\hat{F}(z)$.

Following the procedure of Section 3.3 we apply Theorem 2.4.3 to obtain as necessary conditions for a saddle-value of ϕ that a system of inequalities involving the marginal productivities and prices must be satisfied. Thus world efficiency of production implies the fulfillment of those inequalities.

THEOREM 3.6.1. *If z is world efficient, then the following inequalities are satisfied. (It is understood that derivatives are evaluated at the point (z, q) at which $\phi(z, q)$ is a saddle-value.)*

$$\text{(i)} \qquad \frac{\partial \phi}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} = p_\alpha \frac{\partial {}^{\mathfrak{I}}g^\alpha}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} - q_\beta \leqq 0\,, \qquad \begin{cases} \alpha = 1^*, \cdots, A^*, \\ \beta = 1^*, \cdots, B^*, \\ j = 1, \cdots, J, \end{cases}$$

$$\text{(ii)} \qquad \frac{\partial \phi}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} = p_\alpha \frac{\partial {}^{\mathfrak{I}}g^\alpha}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} - {}^{\mathfrak{I}}q_\beta \leqq 0\,, \qquad \begin{cases} \alpha = 1^*, \cdots, A^*, \\ \beta = 1^{\mathfrak{I}}, \cdots, B^{\mathfrak{I}}, \\ j = 1, \cdots, J, \end{cases}$$

$$\text{(iii)} \qquad \frac{\partial \phi}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} = {}^{\mathfrak{I}}p_\alpha \frac{\partial {}^{\mathfrak{I}}g^\alpha}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} - q_\beta \leqq 0\,, \qquad \begin{cases} \alpha = 1^{\mathfrak{I}}, \cdots, A^{\mathfrak{I}}, \\ \beta = 1^*, \cdots, B^*, \\ j = 1, \cdots, J, \end{cases}$$

$$\text{(iv)} \qquad \frac{\partial \phi}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} = {}^{\mathfrak{I}}p_\alpha \frac{\partial {}^{\mathfrak{I}}g^\alpha}{\partial {}^{(\mathfrak{I})}z_\beta^\alpha} - {}^{\mathfrak{I}}q_\beta \leqq 0\,, \qquad \begin{cases} \alpha = 1^{\mathfrak{I}}, \cdots, A^{\mathfrak{I}}, \\ \beta = 1^{\mathfrak{I}}, \cdots, B^{\mathfrak{I}}, \\ j = 1, \cdots, J, \end{cases}$$

$$\text{(v)} \qquad \sum_{\mathfrak{I}=1}^{J}\left[{}^{\mathfrak{I}}c_\beta - \left(\sum_{\alpha=1^*}^{A^*} {}^{(\mathfrak{I})}z_\beta^\alpha + \sum_{\alpha=1^{\mathfrak{I}}}^{A^{\mathfrak{I}}} {}^{(\mathfrak{I})}z_\beta^\alpha\right)\right] \geqq 0\,, \qquad \beta = 1^*, \cdots, B^*,$$

$$\text{(vi)} \qquad {}^{\mathfrak{I}}c_\beta - \left(\sum_{\alpha=1^*}^{A^*} {}^{(\mathfrak{I})}z_\beta^\alpha + \sum_{\alpha=1^{\mathfrak{I}}}^{A^{\mathfrak{I}}} {}^{(\mathfrak{I})}z_\beta^\alpha\right) \geqq 0\,, \qquad \begin{cases} \beta = 1^{\mathfrak{I}}, \cdots, B^{\mathfrak{I}}, \\ j = 1, \cdots, J. \end{cases}$$

Following the procedure of Section 3.3 we define unspecialized world efficiency.

DEFINITION 3.6.1. *The world efficient point z is called unspecialized if $z > 0$ and $q > 0$, where q is associated with z.*

ASSUMPTION 3.6.1. *The point z is an unspecialized world efficient point.*

Again, in the light of Theorem 2.4.3, it follows that the inequalities (i)–(vi) of Theorem 3.6.1 are all equalities.

As before, we wish to assure the existence of a world efficient point for each $p > 0$ and each vector c of resource endowments.

ASSUMPTION 3.6.2. *There is an unspecialized world efficient point.*

Applying the argument following Theorem 3.3.2 to the present case, the factor price equalization problem is reduced to the problem of considering solutions to the system of equations obtained by replacing the inequalities (i)–(iv) of Theorem 3.6.1 by equalities.

Following the outline of step by step generalization given above, consider first the case in which only equations of type (ii) are present.

ASSUMPTION 3.6.3. $A^1 = B^* = 0$.

The question of finding necessary and sufficient conditions for factor price equalization has been reduced to the question of finding necessary and sufficient conditions that the equations (ii) have unique solutions ${}^jq_i$, with the property

$$ {}^iq_i = {}^jq_i \qquad (i, j = 1, \cdots, J) . $$

Consider equations (ii) for fixed j. As in the earlier case, elimination of ${}^jq_i$ results in the following equations:

$$ (3.6.1) \qquad p_{1^*} \frac{\partial^j g^{1^*}}{\partial^{(j)} z_i^{1^*}} = p_{2^*} \frac{\partial^j g^{2^*}}{\partial^{(j)} z_i^{2^*}} = \cdots = p_{A^*} \frac{\partial^j g^{A^*}}{\partial^{(j)} z_i^{A^*}} \qquad (i = 1^1, \cdots, B^1). $$

These are $B^1(A^* - 1)$ equations in the B^1A^* partial derivatives. We consider two sub-cases, depending on whether the number of products is larger or smaller than the number of resources.

In case $A^* \geqq B^1$, the A^* additional conditions

$$ V^{\alpha^*}\left(\frac{\partial^j g^{\alpha^*}}{\partial^{(j)} z_{1^1}^{\alpha^*}}, \cdots, \frac{\partial^j g^{\alpha^*}}{\partial^{(j)} z_{B^1}^{\alpha^*}}\right) = 0 \qquad (\alpha = 1^*, \cdots, A^*) $$

will be at least enough to determine the partial derivatives, for if $A^* \geqq B^1$,

$$ B^1A^* \leqq B^1A^* - B^1 + A^* . $$

The consistency of the equations is assured by the hypothesis that $\bar{z}$ is world efficient and by a condition on the Jacobian of the functions V^{α^*}. Thus, the proof of Theorem 3.3.2 generalizes in a straightforward manner to yield a sufficient condition for factor price equalization in the present case.

THEOREM 3.6.2. *Let V^{α^*}, $\alpha = 1^*, \cdots, A^*$ be continuously differentiable functions of B^1 variables, $A^* \geqq B^1$, having the property that there exists a renumbering of the equations*

$$ V^{\alpha^*}\left(\frac{\partial^j g^{\alpha^*}}{\partial^{(j)} z_{1^1}^{\alpha^*}}, \cdots \frac{\partial^j g^{\alpha^*}}{\partial^{(j)} z_{B^1}^{\alpha^*}}\right) \equiv 0 \qquad \begin{cases} \alpha = 1^*, \cdots, A^*, \\ j = 1, \cdots, J , \end{cases} $$

such that the conditions of Lemma 3.3.1 *are satisfied for the first* B^1 *principal minors, and every minor after the* B^1*-th is zero. Then, the equations* (*ii*) *have a unique solution* ${}^jq_i$ *which is independent of* j.

In other words, the world efficiency of $\bar{z}$ under these conditions implies factor price equalization. For necessary conditions, the proof of Theorem 3.3.3 carries over to establish 3.6.3.

THEOREM 3.6.3. *Let every solution* q_i *of equations* (*ii*) *have the property that* ${}^jq_i$ *is independent of* j (*i.e.*, ${}^iq_i = {}^jq_i$ *for all* $i, j = 1, 2, \cdots, J$). *Then it follows that there exist functions* $V^{\alpha^*}(\ )$ *of* B^1 *variables such that* ${}^jg^{\alpha^*}$ *satisfy*

$$V^{\alpha^*}\left(\frac{\partial {}^jg^{\alpha^*}}{\partial {}^{(j)}z^{\alpha^*}_{i_1^1}}, \cdots, \frac{\partial {}^jg^{\alpha^*}}{\partial {}^{(j)}z^{\alpha^*}_{i_{B^1}^1}}\right) = 0$$

for $\bar{z}$ *unspecialized world efficient.*

In the case $A^* < B^1$ there are too few equations, even with functional dependence among the partial derivatives of the production functions, to determine the ${}^jq_i$ independently of ${}^jc$.

The second step in our program of generalization is to consider the case of mobile resources and immobile products. In this case we replace Assumption 3.6.3 by

ASSUMPTION 3.6.4. $A^* = B^1 = 0$.

In this case only equations (iii) of Theorem 3.6.1 apply (are not identically zero). Elimination of the ${}^jp_i$ gives us the following well-known system of equations:

$$\frac{\dfrac{\partial {}^jg^{\alpha}}{\partial {}^{(j)}z^{\alpha}_{i_{\beta_1}}}}{\dfrac{\partial {}^jg^{\alpha}}{\partial {}^{(j)}z^{\alpha}_{i_{\beta_2}}}} - \frac{q_{i_{\beta_1}}}{q_{i_{\beta_2}}} = 0 \qquad \begin{cases} \alpha = 1^1, \cdots, A^1, \\ \beta = 1^*, \cdots, B^*. \end{cases}$$

For each j there are $A^1(B^* - 1)$ equations and A^1B^* partial derivatives to be determined. Hence A^1 additional equations

$$V^{\alpha}\left(\frac{\partial {}^jg^{\alpha}}{\partial {}^{(j)}z^{\alpha}_{i_1^*}}, \cdots, \frac{\partial {}^jg^{\alpha}}{\partial {}^{(j)}z^{\alpha}_{i_{B^*}^*}}\right) = 0 \qquad (\alpha = 1^1, \cdots, A^1)$$

will always be sufficient in number, for

$$A^1B^* = A^1B^* - A^1 + A^1 .$$

Thus, the preceding argument applies.

We come now to the case of both mobile and immobile products, together with immobile resources.

ASSUMPTION 3.6.5. (1) $A^* \geqq B^*$, (2) $A^I > 0$, (3) $B^* = 0$.

In this case only equations (ii) and (iv) of Theorem 3.6.2 apply. If there is a solution q^I of equations (ii) which satisfies the conditions for factor price equalization, then the solution can be substituted into equations (iv), thus reducing the problem to a case already treated. Thus, this case depends on the solution of equations (ii).

Similarly, the case of immobile products with both mobile and immobile resources is reduced to a previously treated case.

ASSUMPTION 3.6.6. (1) $A^* = 0$, (2) $A^I > 0$, (3) $B^* > 0$, (4) $B^I > 0$.

In this case equations (ii) and (iv) of Theorem 3.6.1 apply. If equations (iii) yield a solution p^* independent of j, then this solution may be substituted in equations (iv) to reduce these to the equations (ii).

Finally, combining this with the case discussed under Assumption 3.4.5 gives the case when all types of commodities are present. It is thus shown that the problem of finding conditions for factor price equalization in the case of all combinations of mobility and immobility is reduced essentially to the classical case in which all products are assumed to be mobile and all resources immobile.

3.7. *Analysis of the conditions for factor price equalization in many dimensions.* The interpretation of the conditions for factor price equalization can be dealt with in terms of two questions. First, what restriction is imposed on a production function by the condition that its marginal productivity functions be functionally dependent? Second, what restriction is imposed by requiring the same condition of functional dependence to be satisfied by two different production functions? We will consider the second question first.

Let f be a (smooth) function of n variables and let the equation

$$V(f_1, f_2, \cdots, f_n) = 0 \tag{3.7.1}$$

hold identically in the arguments of f.

We may consider equation (3.7.1) as a partial differential equation which we know to be satisfied by the function f. Our question is, can this partial differential equation have more than one solution? It is quite clear that if no further conditions are imposed, 3.7.1 could have infinitely many solutions. If in parallel to Assumption 3.4.1 we impose boundary conditions on the problem, e.g., require

$$f(0, x_2, \cdots, x_n) = 0 \qquad \text{for all non-negative } x_2, \cdots, x_n, \tag{3.7.2}$$

we still require V to satisfy certain regularity conditions in order to

insure uniqueness.[23] The regularity conditions V is required to satisfy are, depending to some extent on the nature of the boundary conditions specified, just those required by the factor price equalization theorem, namely non-zero minors of the functional determinant.

While it is clear that there are conditions under which 3.7.1 will have a unique solution, and therefore under which the condition for factor price equalization implies that production functions must be the same in all countries, these boundary conditions do not seem to be intrinsic in the factor price equalization problem. Which of the available possibilities should be imposed, including the possibility of no condition, depends on the economic meaning of the conditions; equation 3.7.2 asserts the existence of an indispensable input, e.g., labor, without which production is impossible.

We turn now to our first question, namely what does the condition that marginal productivity functions be functionally dependent imply about the production function? In the many dimensional case, in contrast to that of two inputs, our conditions do not seem to imply homogeneity of the production function. I have, at least, been unable to obtain such a result. It is perhaps of interest to point up the importance of regularity conditions in this problem. For example, the function

$$f(x, y, z) = xe^{y+z}$$

has the property that its partial derivatives are functionally dependent, for

$$f_y - f_z = 0 ,$$

and

$$f(0, y, z) = 0 \qquad \text{for all } y \text{ and } z,$$

but f is obviously not homogeneous. The equation of functional dependence fails to satisfy regularity conditions insuring solvability for f_z on the plane $x = 0$. This example also points up the importance of regularity conditions in guaranteeing uniqueness, for

$$g(x, y, z) = x(y + z)$$

also satisfies

$$g_y - g_z = 0 \qquad \text{with } g(0, y, z) = 0 .$$

[23] See Dorothy L. Bernstein, *Existence Theorems in Partial Differential Equations*, Annals of Mathematics, Study No. 23, (Princeton, New Jersey: Princeton University Press 1950). See in particular Theorem 13.1, p. 49 and subsequent remarks.

3.8. *Discussions of Assumptions* 3.3.2, 3.6.2. The effect of Assumption 3.3.2 is to restrict attention to unspecialized world efficient points, i.e., to those in which every country produces every commodity and uses every resource to the limit of availability.

Consider first the matter of resources. Recall that a necessary condition for factor price equalization is that the same resources be scarce in every country (see the remark at the end of Section 3.1). We can in any case ignore the set of resources which are superabundant at a particular world efficient point, since the solutions of the relevant equations are unaffected by the restrictions corresponding to those resources. So, at least with regard to resources, Assumption 3.3.2 and its later counterpart, Assumption 3.6.2, constitute no real restriction of generality.

With regard to products, the assumption does impose a real restriction. Returning our attention to the full system of inequalities (i) through (iv) of Theorem 3.6.1, suppose that the given world efficient point is not unspecialized, i.e., that each country produces some commodities but not all. We direct our attention to those commodities which every country produces, and to the corresponding sub-system of equations. Clearly the foregoing analysis applies to this sub-system, and hence to the whole. If the international economy is sufficiently connected through production of mobile commodities, then there will be a sufficient number of prices which are the same in all countries, and a sufficient number of equations in each country so that the preceding analysis applies. In this light, the role of Assumption 3.6.3 has been merely to provide a sharp focus for the discussion.

3.9. *The actual economy.* Attempts to apply factor price equalization theorems to an actual economy will face two questions deserving brief comment here:

First, under what circumstances is, e.g., a bricklayer in London to be regarded the same factor as a bricklayer in New York? We may make a "thought experiment" in which the two bricklayers are in turn combined with the same specific combination of other resources. If the two resulting relations between output and input are the same, the two factors may be regarded as one. Arguing from the "thought experiment" that it is meaningful to speak of the same resource in different countries, the problem of identifying resources is not essentially different from many similar problems of classification encountered in applications of economic theory.

Second, granting that it is meaningful to ask whether factor prices are equalized, does not the dependence of the factor price equalization

theorem on the relative number of mobile and immobile commodities make the results arbitrary in light of the fact that commodities can be classified arbitrarily? To this we point out that it is (a) possible to classify commodities arbitrarily and (b) possible to classify commodities so as to give the production function any desired form, but that (c) it is not in general possible to specify arbitrarily both the form of the production function and the classification of inputs and outputs and still describe adequately given physical production processes. Because the factor price equalization theorems depend on both the properties of the production function and the classification of commodities, we may expect that there is more to the results than mere tricks of counting.

Purdue University, U.S.A.

INTERNATIONAL ECONOMIC REVIEW
Vol. 2, No. 1, January, 1961

ON PURE AND MIXED STATISTICAL ESTIMATION IN ECONOMICS*

BY H. THEIL AND A. S. GOLDBERGER

1. INTRODUCTION

A BASIC PROBLEM in the statistical estimation of economic relations is the discrepancy between the statistician's claim when he formulates his maintained hypothesis and the procedure which he actually follows. He claims that he includes all his relevant *a priori* information in the maintained hypothesis, and it is on the maintenance of this hypothesis that his probability statements are based. Yet it is not uncommon to find statisticians discarding their maintained hypotheses after the estimation has been undertaken.

For example, the investigator will accept an economic theory which explains the variation in some variable linearly in terms of the variations in a fixed set of other variables. The linear relation is assumed to be subject to additive disturbances with, say, zero means, constant variances, and zero covariances. All this makes up his maintained hypothesis. He then carries out an appropriate computation to provide desirable estimates of the parameters of the linear relation. He finds that the estimated income elasticity of some commodity (which is unlikely to be inferior) is negative. He does not accept this negative estimate but rather attributes the result to the incorrectness of his previously maintained hypothesis and perhaps decides to change the set of explanatory variables. It is well known, but also well ignored, that exact probability statements can no longer be made if the maintained hypothesis is thus rejected in the light of the evidence.

Recollection of our own past activities prevents us from denouncing too violently such an investigator, but we agree that a more consistent approach should be sought. The difficulty seems to be that the investigator has *a priori* knowledge which he can not conveniently incorporate in the maintained hypothesis and which he therefore omits. In particular, the *a priori* knowledge that a regression slope is positive (as in the example of the preceding paragraph) can not readily be built into a maintained hypothesis by conventional methods. This kind of *a priori* knowledge, however, is precisely the major source of rejections of "maintained hypotheses"; it seems clear that it is logically more consistent to incorporate such knowledge in the maintained hypothesis

* Manuscript received April 9, 1960, revised August 2, 1960.

right at the beginning than to exclude it from the hypothesis and reject it afterwards when the results contradict the omitted knowledge.

In this paper we propose a method for handling *a priori* inequalities on coefficients, or on linear combinations of coefficients, as an integral part of the maintained hypothesis. This method may be regarded as an extension of a method for efficient use of extraneous *statistical* information which will first be discussed in detail.

The order of discussion is as follows: In Section 2 we consider the problem of the efficient use of extraneous statistical information in linear regression; Section 3 is devoted to the extension of this procedure to the case in which this extraneous information is of an *a priori*, rather than a statistical, type; Section 4 contains an illustrative example.

2. EFFICIENT USE OF EXTRANEOUS STATISTICAL INFORMATION

2.1. *Application of generalized least squares.* Let us suppose that we have a sample of observations specified to satisfy the linear relationship

$$y = X\beta + u , \tag{2.1}$$

where y is the $T \times 1$ vector of observations on the dependent variable, X the $T \times A$ matrix of observations on the explanatory variables with rank A, β the $A \times 1$ vector of coefficients, and u the $T \times 1$ vector of disturbances, such that

$$Eu = 0 , \qquad E(uu') = \Omega , \tag{2.2}$$

with Ω non-singular. It will further be assumed that the elements of X are nonstochastic.

Furthermore, we assume that additional information on the coefficients β is available. Here, in Section 2, it will be assumed that this knowledge is of the statistical type (it may, for example, come from previous statistical investigations), and that it consists of estimates of β-components or of linear combinations of these. In any event, we suppose that it is possible to represent this extraneous information by

$$r = R\beta + v , \tag{2.3}$$

where r is a $G \times 1$ vector, R is a given $G \times A$ matrix, β is the $A \times 1$ coefficient vector of (2.1), and v is a $G \times 1$ vector of disturbances such that

$$Ev = 0 , \qquad E(vv') = \Psi , \tag{2.4}$$

with Ψ non-singular.

The additional information might, for example, consist of b_1 and b_2, unbiased estimates of the first two elements of β, along with their variances, σ_{11} and σ_{22}, and their covariance, σ_{12}. In that case we would write

$$(2.5) \qquad r = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{bmatrix}, \quad \Psi = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}.$$

The sample and the independent extraneous information may be combined from the start by writing

$$(2.6) \qquad \begin{bmatrix} y \\ r \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix}\beta + \begin{bmatrix} u \\ v \end{bmatrix}; \quad E\begin{bmatrix} u \\ v \end{bmatrix} = 0; \quad E\left(\begin{bmatrix} u \\ v \end{bmatrix}[u' \;\; v']\right) = \begin{bmatrix} \Omega & 0 \\ 0 & \Psi \end{bmatrix}.$$

An application of Aitken's generalized least-squares procedure leads to estimating β as $\hat{\beta}$, where

$$(2.7) \qquad \hat{\beta} = \left([X' \;\; R']\begin{bmatrix} \Omega & 0 \\ 0 & \Psi \end{bmatrix}^{-1}\begin{bmatrix} X \\ R \end{bmatrix}\right)^{-1}[X' \;\; R']\begin{bmatrix} \Omega & 0 \\ 0 & \Psi \end{bmatrix}^{-1}\begin{bmatrix} y \\ r \end{bmatrix},$$

that is,

$$(2.8) \qquad \hat{\beta} = [X'\Omega^{-1}X + R'\Psi^{-1}R]^{-1}[X'\Omega^{-1}y + R'\Psi^{-1}r].$$

This estimate has the property of being a best linear unbiased estimate, with "best" referring to both the extraneous and the sample information. The variance-covariance matrix of the estimate is given by

$$(2.9) \qquad V(\hat{\beta}) = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] = [X'\Omega^{-1}X + R'\Psi^{-1}R]^{-1}.$$

The estimating formula will, of course, be simplified if the u's are mutually independent with a common variance. In that case, $\Omega = \sigma^2 I$, so that

$$(2.10) \qquad \hat{\beta} = \left[\frac{1}{\sigma^2}X'X + R'\Psi^{-1}R\right]^{-1}\left[\frac{1}{\sigma^2}X'y + R'\Psi^{-1}r\right].$$

If, in addition, there is only a single piece of extraneous information, say b_1, an unbiased estimate of the first element of β with variance σ_*^2, then

$$R'\Psi^{-1}R = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\frac{1}{\sigma_*^2}[1 \;\; 0 \;\; \cdots \;\; 0] = \frac{1}{\sigma_*^2}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

and

$$R'\Psi^{-1}r = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\frac{1}{\sigma_*^2}b_1 = \frac{1}{\sigma_*^2}\begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Hence,

$$\hat{\beta} = \left(X'X + \frac{\sigma^2}{\sigma_*^2} \begin{bmatrix} 1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{bmatrix} \right)^{-1} \left(X'y + \frac{\sigma^2}{\sigma_*^2} \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right). \tag{2.11}$$

In this case it will be seen that the normal equations defining $\hat{\beta}$ differ from those that would be used in the absence of extraneous information only in the upper left-hand element of the coefficient matrix and in the top element of the right-hand vector.

This simple result illustrates at the same time a difficulty in applying the method proposed, viz., the fact that knowledge of Ω and Ψ is required (in the case of (2.11), knowledge of σ^2 and σ_*^2, or at least of their ratio). It can be shown, however, that if we commit small errors in the specification of Ω and Ψ, the resulting errors at the level of $\hat{\beta}$ tend to be small of a higher order (cf. Van IJzeren [6]). We shall return to this problem in a later publication.

2.2. *Alternative interpretation of the procedure.* The estimate $\hat{\beta}$ of (2.8) was, in fact, proposed several years ago by J. Durbin [2], who was concerned with the case where the extraneous information comprised an unbiased estimate of a subvector of β (along with the corresponding variances and covariances) obtained from another sample. Durbin's approach was to blend this extraneous estimate with the "internal" least-squares estimate $\tilde{\beta} = (X'X)^{-1}X'y$ in an efficient fashion. This led to the estimator of (2.8). It should also be noted that J. R. N. Stone [3], pp. 303–305, has given a maximum-likelihood interpretation of the same estimator in a special case.

Durbin's approach is suggestive. Clearly, if the extraneous information does not differ from the internal estimates, then the efficient estimate will be the internal one, $\hat{\beta} = \tilde{\beta}$. Further, intuition suggests that the difference between $\hat{\beta}$ and $\tilde{\beta}$ will depend in a systematic way upon the degree to which the internal estimate $\tilde{\beta}$ differs from the "mean extraneous information," i.e., upon the difference between r and $R\tilde{\beta}$. This can be analyzed in the following way: The matrix to be inverted in (2.8) is

$$X'\Omega^{-1}X + R'\Psi^{-1}R = X'\Omega^{-1}X[I + (X'\Omega^{-1}X)^{-1}R'\Psi^{-1}R] .$$

Suppose that the second matrix between square brackets is sufficiently "small" in the sense that its latent roots are all within the unit circle; in qualitative terms this means that the internal X information dominates the extraneous R information. We can now apply a power series expansion for the inverse, and it follows that:

$$(X'\Omega^{-1}X + R'\Psi^{-1}R)^{-1} = [I + (X'\Omega^{-1}X)^{-1}(R'\Psi^{-1}R)]^{-1}(X'\Omega^{-1}X)^{-1}$$
$$\approx (X'\Omega^{-1}X)^{-1} - (X'\Omega^{-1}X)^{-1}(R'\Psi^{-1}R)(X'\Omega^{-1}X)^{-1},$$

where second-order terms are neglected. Combining this with (2.8) we obtain

$$\begin{aligned}(2.12)\qquad \hat{\beta} &\approx (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y + (X'\Omega^{-1}X)^{-1}R'\Psi^{-1}r \\ &\quad - (X'\Omega^{-1}X)^{-1}(R'\Psi^{-1}R)(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \\ &\approx \tilde{\beta} + (X'\Omega^{-1}X)R'\Psi^{-1}(r - R\tilde{\beta}),\end{aligned}$$

where second-order terms are again neglected. Thus, to a first-order approximation at least, the difference between $\hat{\beta}$ and $\tilde{\beta}$ is a homogeneous linear function of the extent to which $\tilde{\beta}$ fails to satisfy the "mean extraneous information."

2.3. *Extension to simultaneous equation estimation.* The technique developed in Section 2.1 may be extended to the estimation of an equation which is embodied in a simultaneous system. That is, we now suppose that our sample of observations satisfies the linear relationship

$$(2.13)\qquad y = Y\gamma + X_1\beta + u = [Y \;\; X_1]\begin{bmatrix}\gamma \\ \beta\end{bmatrix} + u,$$

where y is the $T \times 1$ vector of observations on the "left-hand" endogenous variable, Y the $T \times m$ matrix of observations on the other endogenous variables present in the equation, γ the $m \times 1$ vector of their coefficients, X_1 the $T \times l$ matrix of observations on the predetermined variables present in the equation, β the $l \times 1$ vector of their coefficients, and u the $T \times 1$ vector of disturbances, such that,

$$(2.14)\qquad Eu = 0, \qquad E(uu') = \sigma^2 I.^1$$

We shall apply the two-stage least-squares procedure. This can be formulated in a variety of ways, the following being the most convenient one for our problem. Multiply (2.13) through by X', where X is the $T \times \Lambda$ matrix of observations on all the predetermined variables in the system. This yields (cf. Theil [4], pp. 337–38),

$$(2.15)\qquad X'y = X'Y\gamma + X'X_1\beta + u^* = [X'Y \;\; X'X_1]\begin{bmatrix}\gamma \\ \beta\end{bmatrix} + u^*,$$

where $u^* = X'u$, and

$$(2.16)\qquad Eu^* = 0; \qquad E(u^*u^{*\prime}) = \sigma^2(X'X).$$

[1] In addition, we make the usual assumption that the disturbances of the other structural equations are also uncorrelated over time and have constant variances and "contemporaneous" covariances. A generalization to the temporally correlated case is also possible (cf. Theil [4], pp. 343–45), but is not pursued here.

The application of generalized least-squares to (2.15) gives

$$\begin{bmatrix} Y'X \\ X_1'X \end{bmatrix}(\sigma^2 X'X)^{-1}X'y = \begin{bmatrix} Y'X \\ X_1'X \end{bmatrix}(\sigma^2 X'X)^{-1}[X'Y \quad X'X_1]\begin{pmatrix} c \\ b \end{pmatrix}, \tag{2.17}$$

where c, b is the two-stage least-squares estimator of γ, β. The estimation equations can be simplified as follows:

$$\begin{bmatrix} Y'X(X'X)^{-1}X'y \\ X_1'y \end{bmatrix} = \begin{bmatrix} Y'X(X'X)^{-1}X'Y & Y'X_1 \\ X_1'Y & X_1'X_1 \end{bmatrix}\begin{pmatrix} c \\ b \end{pmatrix}, \tag{2.18}$$

and the asymptotic covariance matrix is

$$\lim_{T\to\infty} TV\begin{pmatrix} c \\ b \end{pmatrix} = \sigma^2 \operatorname*{plim}_{T\to\infty} T\begin{bmatrix} Y'X(X'X)^{-1}X'Y & Y'X_1 \\ X_1'Y & X_1'X_1 \end{bmatrix}^{-1}. \tag{2.19}$$

The procedure presupposes that $A \geqq m + l$; if the opposite is true, the square matrix on the right of (2.18) is singular, and the coefficients of (2.13) are not identifiable.

The extraneous information is now written as

$$r = R_1\gamma + R_2\beta + v, \tag{2.20}$$

where R_1 and R_2 are matrices of known coefficients and v a vector of random elements satisfying (2.4). We can combine the sample and extraneous information by writing

$$\begin{bmatrix} X'y \\ r \end{bmatrix} = \begin{bmatrix} X'Y & X'X_1 \\ R_1 & R_2 \end{bmatrix}\begin{bmatrix} \gamma \\ \beta \end{bmatrix} + \begin{bmatrix} u^* \\ v \end{bmatrix}, \tag{2.21}$$

where

$$E\begin{bmatrix} u^* \\ v \end{bmatrix} = 0, \quad \text{and} \quad \left(E\begin{bmatrix} u^* \\ v \end{bmatrix}[u^{*\prime} \quad v']\right) = \begin{bmatrix} \sigma^2 X'X & 0 \\ 0 & \Psi \end{bmatrix}. \tag{2.22}$$

Hence the left-hand side of (2.17) becomes

$$\begin{bmatrix} Y'X & R_1' \\ X_1'X & R_2' \end{bmatrix}\begin{bmatrix} \sigma^2 X'X & 0 \\ 0 & \Psi \end{bmatrix}^{-1}\begin{bmatrix} X'y \\ r \end{bmatrix},$$

while the right-hand side consists of a square matrix,

$$\begin{bmatrix} Y'X & R_1' \\ X_1'X & R_2' \end{bmatrix}\begin{bmatrix} \sigma^2 X'X & 0 \\ 0 & \Psi \end{bmatrix}^{-1}\begin{bmatrix} X'Y & X'X_1 \\ R_1 & R_2 \end{bmatrix},$$

postmultiplied by the vector of estimates, $\hat{\gamma}, \hat{\beta}$. Simplifying and writing $\Psi_* = (1/\sigma^2)\Psi$, we obtain

$$\begin{aligned}\begin{pmatrix} \hat{\gamma} \\ \hat{\beta} \end{pmatrix} = &\begin{bmatrix} Y'X(X'X)^{-1}X'Y + R_1'\Psi_*^{-1}R_1 & Y'X_1 + R_1'\Psi_*^{-1}R_2 \\ X_1'Y + R_2'\Psi_*^{-1}R_1 & X'X_1 + R_2'\Psi_*^{-1}R_2 \end{bmatrix}^{-1} \\ &\times \begin{bmatrix} Y'X(X'X)^{-1}X'y + R_1'\Psi_*^{-1}r \\ X_1'y + R_2'\Psi_*^{-1}r \end{bmatrix},\end{aligned} \tag{2.23}$$

while the asymptotic covariance matrix is σ^2 multiplied by the probability limit of T times the inverse matrix of (2.23).

2.4. *Balanced use of statistical information.* There may be some asymmetry involved in the procedure of Section 2.1. Suppose there are two sets of sample observations on y and the explanatory variables. We might estimate parameters from the first sample and use the estimates as extraneous information for efficient estimation in the second sample. But the roles of the two samples might well be reversed. A more balanced approach would put the two samples on an equal footing for the estimation of parameters. For this approach, a natural extension of the procedure of Section 2.1 suggests itself.

Suppose there are two samples A and B, each of which is described by a separate relation of the form (2.1). Typically, the relation describing sample A will have some parameters in common with the relationship describing sample B and some others distinctive to it, while the relationship describing sample B will have some distinctive parameters besides those it shares with the A relationship. This could be the case in household demand studies. That is, y might be per capita quantity of meat, sample A consisting of time series observations on y, per capita income, and price of meat; while sample B might consist of cross section observations on y, per capita income, and family size. The two relationships could share the same income coefficient.

For such a situation, all the observations would be described by the relationship

$$\begin{bmatrix} y_A \\ y_B \end{bmatrix} = \begin{bmatrix} X_{A0} & X_{A1} & 0 \\ X_{B0} & 0 & X_{B2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_A \\ u_B \end{bmatrix}, \tag{2.24}$$

where the subscripts A and B denote data from samples A and B respectively, and the subscripts $0, 1, 2$ refer to parameters (and corresponding variables) which are, respectively, common to both samples, distinctive to sample A, and distinctive to sample B.

If we assume

$$E\begin{bmatrix} u_A \\ u_B \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad E\left(\begin{bmatrix} u_A \\ u_B \end{bmatrix}[u_A' \quad u_B']\right) = \begin{bmatrix} \Omega_A & 0 \\ 0 & \Omega_B \end{bmatrix}, \tag{2.25}$$

and apply the generalized least-squares procedure, we obtain estimates of the coefficients satisfying the normal equations

(2.26)

$$\begin{bmatrix} X_{A0}' & X_{B0}' \\ X_{A1}' & 0 \\ 0 & X_{B2}' \end{bmatrix} \begin{bmatrix} \Omega_A^{-1} & 0 \\ 0 & \Omega_B^{-1} \end{bmatrix} \begin{bmatrix} y_A \\ y_B \end{bmatrix} = \begin{bmatrix} X_{A0}' & X_{B0}' \\ X_{A1}' & 0 \\ 0 & X_{B2}' \end{bmatrix} \begin{bmatrix} \Omega_A^{-1} & 0 \\ 0 & \Omega_B^{-1} \end{bmatrix} \begin{bmatrix} X_{A0} & X_{A1} & 0 \\ X_{B0} & 0 & X_{B2} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix},$$

that is

$$(2.27)\qquad \begin{bmatrix} X'_{A0}\Omega_A^{-1}y_A + X'_{B0}\Omega_B^{-1}y_B \\ X'_{A1}\Omega_A^{-1}y_A \\ X'_{B2}\Omega_B^{-1}y_B \end{bmatrix} = \begin{bmatrix} X'_{A0}\Omega_A^{-1}X_{A0} + X'_{B0}\Omega_B^{-1}X_{B0} & X'_{A0}\Omega_A^{-1}X_{A1} & X'_{B0}\Omega_B^{-1}X_{B2} \\ X'_{A1}\Omega_A^{-1}X_{A0} & X'_{A1}\Omega_A^{-1}X_{A1} & 0 \\ X'_{B2}\Omega_B^{-1}X_{B0} & 0 & X'_{B2}\Omega_B^{-1}X_{B2} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}.$$

The resulting estimator is best linear unbiased with respect to the two samples combined, if the elements of all X-matrices are non-stochastic; in that case, the covariance matrix is equal to the inverse of the square matrix on the right of (2.27).

3. THE METHOD OF MIXED LINEAR ESTIMATION

Suppose now that the source of extraneous information is of an *a priori* rather than a statistical nature. For example, suppose that the investigator has *a priori* knowledge that a certain coefficient β_1—an income elasticity—almost certainly lies between zero and one, and probably between $\frac{1}{4}$ and $\frac{3}{4}$. Then we propose that he formulate this knowledge in the spirit of Section 2.1 by writing

$$(3.1)\qquad \tfrac{1}{2} = \beta_1 + v\,; \qquad Ev = 0\,; \qquad Ev^2 = 1/16.$$

Values of β_1 outside the range between zero and one would then be outside the "two-sigma" range $\frac{1}{2} \pm 2\sqrt{1/16}$; this gives some indication of the virtual impossibility of such values. Once the pieces of *a priori* information are expressed in this way, the procedure of Section 2.1 is applicable to obtain "best linear estimates." Here "best" refers not only to the observations but also to *a priori* knowledge. We shall refer to this procedure as "mixed linear estimation" in view of the mixed character of the information on which it is based; for the "ordinary" procedure, which is based on observational data alone, we shall use the term "pure linear estimation."

Several pieces of *a priori* information may, of course, be handled simultaneously. Knowledge about linear combinations of coefficients may also be handled similarly; for example, the sum of all price and income elasticities of a demand equation must be approximately zero because there cannot be "too much" money illusion. Clearly, the narrower the interval within which a coefficient almost certainly lies, the more useful is the incorporation of this knowledge. In the extreme situation, where the interval shrinks to zero, we have the case of exact *a priori* restrictions. In that instance, the *a priori* information might be used to eliminate part of the coefficient vector to be

estimated, after which simple least-squares would be applied to the "reduced" equation (cf., e.g., Theil [4] p. 331). As the interval widens, that is, as the *a priori* information becomes hazier, it plays less of a role in the estimation. Thus, starting from any given Ψ matrix, Ψ_0 say, we put $\Psi = k\Psi_0$; then as k approaches infinity, Ψ^{-1} approaches 0, so that, as intuition suggests and (2.8) confirms, the best estimate is simply the internal least-squares estimate; in other words, the *a priori* "information" is useless.

With the present approach, mere knowledge that a coefficient is positive (or that it is negative) would be of no use since the relevant interval—and hence Ψ—would be infinite. It would, however, seem to be typical for the investigator to be able to set a finite upper bound which the absolute value of the parameter almost certainly does not exceed.

It is appropriate to examine the expectations of (3.1) in somewhat more detail. The probability concept underlying these expectations is evidently of the subjectivist or "personalist" type, where the person involved is the research worker who formulates the *a priori* point estimate as well as its "sampling variance." Our idea is that such a person carries out a large number of mixed regression analyses in the course of his lifetime, so that he produces a similarly large number of errors (the differences between the *a priori* estimates and the corresponding parameters). Dividing these errors by the corresponding square roots of their variances, we obtain standardized errors which are—in the light of this subjectivist probability idea—random drawings from a parent with zero mean and unit variance. It is worth-while in this connection to add that the unbiasedness condition (like $Ev = 0$ in (3.1)) is not required for all purposes; this is important in view of the fact that in several cases it will be easier to formulate probable limits for a coefficient than a central tendency. Let us take (3.1); if we accept the weaker condition $Ev^2 = 1/16$ but Ev not necessarily $= 0$, then the resulting mixed estimator of β is of course no longer unbiased, but its second-order moment matrix is still equal to the original covariance matrix, as can be easily verified.

4. NUMERICAL ILLUSTRATION

For illustrative purposes we apply our technique to the estimation of the supply equation of Tintner's model of the American meat market [5], pp. 176–179. It has been shown by Begeer [1] that the ordinary two-stage least-squares estimates do not give the "right" signs for all of the coefficients.[2] Under the standard procedure, therefore, this

[2] The same is true, incidentally, of the "ordinary" single-stage least-squares estimates.

relation would be discarded or modified, i.e., the maintained hypothesis would be rejected. We investigate the effect of including the *a priori* sign criteria right from the start.

The model consists of these two equations:

$$y_1 = \gamma y_2 + \beta_1 x_1 + \beta_2 x_2 + u \qquad \text{(supply)}, \tag{4.1}$$

$$y_1 = \gamma' y_2 + \beta_3 x_3 + u' \qquad \text{(demand)}, \tag{4.2}$$

where y_1 = per capita consumption (pounds per year); y_2 = retail price (index, 1935–39 = 100); x_1 = cost of processing meat (index, 1935–39 = 100); x_2 = cost of producing agricultural products (index, 1935–39 = 100); x_3 = per capita real disposable income (dollars per year).

The sample consists of annual data, 1919–1941, measured in deviations from the means. Table 1 presents the sums of squares and products.

TABLE 1

SUMS OF SQUARES AND PRODUCTS OF DEVIATIONS FROM THE MEANS

	y_1	y_2	x_1	x_2	x_3
y_1	1369.54	−352.55	−536.48	983.86	3671.91
y_2		1581.49	850.33	1235.76	8354.59
x_1			2534.80	730.78	3611.72
x_2				2626.99	12204.77
x_3					83433.65

We assume temporal independence of the error terms so that the internal two-stage least-squares estimates of the supply equation are obtained according to (2.18) as

$$\begin{bmatrix} \tilde{\gamma} \\ \tilde{\beta}_1 \\ \tilde{\beta}_2 \end{bmatrix} = \begin{bmatrix} 937.93 & 850.33 & 1235.76 \\ 850.33 & 2534.80 & 730.78 \\ 1235.76 & 730.78 & 2626.99 \end{bmatrix}^{-1} \begin{bmatrix} 207.64 \\ -536.48 \\ 983.86 \end{bmatrix} = \begin{bmatrix} -0.3210 \\ -0.2777 \\ 0.6028 \end{bmatrix}. \tag{4.3}$$

In the terminology of Section 3, these are pure estimates. It will be seen that the sign of $\tilde{\gamma}$ is negative, contrary to accepted theory. The variance of the error term u is estimated by dividing the sum of squared residuals by the degrees of freedom; this gives $\hat{\sigma}^2 = 21.13$.

Now let us suppose that extraneous information on the price elasticity of supply had been incorporated from the start. General considerations concerning the nature of agricultural production imply that this price elasticity lies between zero and one. We formulate this by saying that we have an *a priori* point estimate of 0.5 for the price elasticity and that its standard deviation is 0.2. This means that we consider a price elasticity outside the interval between zero and one

to be very implausible. This condition on the elasticity may be translated into a condition on the slope given the sample means $\bar{y}_1 = 166.19$ and $\bar{y}_2 = 92.34$. We take the slope equal to 0.9 ($= 0.5 \times 166.19/92.34$) plus or minus a random variable with zero expectation and standard deviation 0.36 ($= 0.2 \times 166.19/92.34$).

In the notation of (2.20) we have

$$(4.4) \quad r = 0.9\,; \quad R_1 = [1]\,; \quad R_2 = [0 \;\; 0]\,; \quad E(vv') = (0.36)^2 = 0.1296\,.$$

We take σ^2 as its internal two-stage least-squares estimate, viz., 21.13, so that Ψ_* of (2.23) is 0.006133 ($= 0.1296/21.13$). Then we see

$$\Psi_*^{-1} = 1/0.006133 = 163.07\,; \qquad R_1'\Psi^{-1}R_2 = [0 \;\; 0]\,;$$

$$(4.5) \quad R_1'\Psi^{-1}R_1 = 163.07\,; \qquad R_2'\Psi^{-1}R_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = R_2'\Psi_*^{-1}r\,;$$

$$R_1'\Psi^{-1}r = (163.07)(0.9) = 146.76\,; \qquad R_2'\Psi^{-1}R_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Application of (2.23) provides the mixed estimates

$$(4.6) \quad \begin{bmatrix} \hat{\gamma} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 1101.00 & 850.33 & 1235.76 \\ 850.33 & 2534.80 & 730.78 \\ 1235.76 & 730.78 & 2626.99 \end{bmatrix}^{-1} \begin{bmatrix} 354.40 \\ -536.48 \\ 983.86 \end{bmatrix} = \begin{bmatrix} 0.1650 \\ -0.3834 \\ 0.4035 \end{bmatrix},$$

in which $\hat{\gamma}$ has the positive sign. The asymptotic sampling variances may be obtained by multiplying each of the diagonal elements of the inverse matrix of (4.6) by σ^2. A new estimate of σ^2 is available, viz., the sum of squared residuals for the sample observations *using the coefficients estimated in* (4.6), divided by the degrees of freedom.[3] This estimate is 50.25, and it yields as asymptotic standard errors of the respective coefficients 0.35, 0.17, and 0.20.

In view of the fact that the estimate of σ^2 which entered the mixed estimation is not the same as that which comes out of it, an iterative procedure is suggested. That is, we take σ^2 as 50.25, so that Ψ_* is 0.00258 ($= 0.1296/50.25$), and adapt (4.5) and (4.6) accordingly to get a new set of estimates of the coefficients. The resulting residual variance provides the basis for a new iteration. In Table 2 we present the results of several of these iterations. The process appears to converge, and we stop at the fifth iteration; at this stage the coefficients are stable to two decimal places. For purposes of comparison we include the internal two-stage least-squares estimates, and also those based on exact *a priori* information as described in the

[3] We continue to take the degrees of freedom as 19, equal to the number of sample observations, 23, minus the number of coefficients being estimated, 4, (including the implicit constant term).

second paragraph of Section 3. That is, we fix γ at 0.9 and apply simple least-squares to the reduced equation in which $y_1 - \gamma y_2$ is the variable to be explained. Standard errors are given in parentheses below the coefficients. We propose to use an iterative procedure in the (virtually inevitable) case in which σ^2 must be estimated from the data.

TABLE 2

ALTERNATIVE ESTIMATES OF THE MEAT SUPPLY EQUATION USING A PRIORI INFORMATION ON PRICE ELASTICITY

Coefficient	Pure	Mixed						Exact Information
		Initial	Iterations					
			1	2	3	4	5	
γ	−0.321 (0.29)	0.165 (0.35)	0.425 (0.34)	0.535 (0.33)	0.573 (0.32)	0.585 (0.31)	0.588 (0.31)	0.900
β_1	−0.278 (0.11)	−0.383 (0.17)	−0.440 (0.19)	−0.464 (0.21)	−0.472 (0.21)	−0.475 (0.21)	−0.475 (0.21)	−0.543 (0.24)
β_2	0.603 (0.15)	0.404 (0.20)	0.297 (0.23)	0.252 (0.23)	0.236 (0.23)	0.231 (0.23)	0.230 (0.23)	0.102 (0.23)
σ^2	21.13	50.25	74.97	87.30	91.79	93.23	93.68	129.57

The economist will have some doubts about our estimate of β_2, which is, after all, the coefficient of a cost of production variable. There are good *a priori* grounds for believing that this coefficient should be negative, though not less than −1. At the risk of being guilty once again of the procedure against which we have argued in Section 1, we have carried out another set of computations. This time we incorporate *a priori* knowledge on the signs of all three coefficients from the start, specifying that the elasticities of y_1 with respect to y_2, x_1, and x_2 are 0.5, −0.5, and −0.5, respectively, each subject to a standard error of 0.2. When this information on the elasticities is translated into information on the linear coefficients (using the further sample means $\bar{x}_1 = 88.42$ and $\bar{x}_2 = 102.22$), we have in the notation of (2.20)

$$r = \begin{bmatrix} 0.90 \\ -0.94 \\ -0.81 \end{bmatrix}; \quad R_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad R_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}; \tag{4.7}$$

$$E(vv') = \begin{bmatrix} 0.1296 & 0 & 0 \\ 0 & 0.1414 & 0 \\ 0 & 0 & 0.1056 \end{bmatrix}.$$

We again take σ^2 as its internal two-stage least-squares estimate, viz., 21.13, which in the notation of (2.23) provides

$$\Psi_*^{-1} = \begin{bmatrix} 163.07 & 0 & 0 \\ 0 & 149.46 & 0 \\ 0 & 0 & 200.13 \end{bmatrix}; \quad R_1'\Psi_*^{-1}R_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

$$(4.8) \quad R_1'\Psi_*^{-1}R_1 = 163.07; \qquad R_1'\Psi_*^{-1}r = \begin{bmatrix} -140.49 \\ -162.10 \end{bmatrix};$$

$$R_1'\Psi_*^{-1}r = 146.76; \qquad R_1'\Psi_*^{-1}R_2 = \begin{bmatrix} 149.46 & 0 \\ 0 & 200.13 \end{bmatrix};$$

$$R_1'\Psi_*^{-1}R_2 = [0 \quad 0].$$

Application of (2.23) provides the mixed estimates

$$(4.9) \quad \begin{bmatrix} \hat{\gamma} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 1101.00 & 850.33 & 1235.76 \\ 850.33 & 2684.26 & 730.78 \\ 1235.76 & 730.78 & 2827.12 \end{bmatrix}^{-1} \begin{bmatrix} 354.40 \\ -676.97 \\ 821.76 \end{bmatrix} = \begin{bmatrix} 0.4118 \\ -0.4440 \\ 0.2255 \end{bmatrix}.$$

The corresponding new estimate of σ^2 is 74.39, which provides the basis for computing standard errors and also for an iterative process, the results of which are presented in Table 3, which has the same format as Table 2. It will be seen that this iteration converges to estimates with the proper sign.

TABLE 3

ALTERNATIVE ESTIMATES OF THE MEAT SUPPLY EQUATION USING A PRIORI INFORMATION ON ALL ELASTICITIES

Coefficient	Pure	Mixed						Exact Information
		Initial	Iterations					
			1	2	3	4	5	
γ	−0.321 (0.29)	0.412 (0.41)	0.849 (0.38)	0.965 (0.33)	0.987 (0.31)	0.991 (0.31)	0.992 (0.31)	0.900
β_1	−0.278 (0.11)	−0.444 (0.19)	−0.556 (0.23)	−0.603 (0.23)	−0.616 (0.22)	−0.619 (0.22)	−0.620 (0.22)	−0.940
β_2	0.603 (0.15)	0.225 (0.23)	−0.068 (0.24)	−0.205 (0.23)	−0.243 (0.22)	−0.251 (0.22)	−0.253 (0.22)	−0.810
σ^2	21.13	74.39	134.11	158.58	164.77	166.14	166.44	300.32

Netherlands School of Economics, The Netherlands and University of Wisconsin, U.S.A.

REFERENCES

[1] BEGEER, W., "Simultaneous-Equations Estimation: Two-Stage Least-Squares, Its Application, and Its Geometry," Report 5703 of the Econometric Institute of the Netherlands School of Economics (1957).

[2] DURBIN, J., "A Note on Regression When There Is Extraneous Information about One of the Coefficients," *Journal of the American Statistical Association*, Vol. 48 (1953), pp. 799-808.

[3] STONE, J. R. N., *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom*, 1920–1938, Vol. I. Cambridge: Cambridge University Press, 1954.

[4] THEIL, H., *Economic Forecasts and Policy*. Amsterdam: North-Holland Publishing Co., 1958.

[5] TINTNER, G., *Econometrics*. New York: John Wiley and Sons, 1952.

[6] VAN IJZEREN, J., "De theoretische zijde van de methode der kleinste kwadraten," *Statistica Neerlandica*, Vol. 8 (1954), pp. 21–45.

INTERNATIONAL ECONOMIC REVIEW
Vol. 2, No. 1, January, 1961

A SCHEMA FOR THE SUPPLY THEORY OF MONEY*

BY KARL BRUNNER[1]

1. INTRODUCTION

THE KEYNESIAN REVOLUTION substituted a new structure for the velocity approach in monetary analysis. This substitution, however, did not attenuate the concept of a (stock) demand function for money. On the contrary, the analytical and empirical issues raised with the emergence of the Keynesian theory as an alternative explanation of monetary (i.e., aggregate) phenomena gave a new impetus to the logical clarification and empirical specification of this demand function.

The supply theory of money, on the other hand, remained uncultivated. Most of the literature presents us with an accumulation of analytical fragments which form at best only building blocks of an empirically significant theory. The analysis of the money supply has still to bridge some major gaps between the conception of a vague idea concerning the structure of a subject matter and the formulation of a testable hypothesis. With one exception, the only coherent schemata available lack an essential feature of hypothesis-construction, viz., the specification of semantic rules or of behavior properties for the parameters. As a result, these schemata must be interpreted either as calculi with "economic names" or structures to be dealt with according to implicit semantic rules. They have no empirical content and cannot be used in a scientific explanation of observable money stock patterns[2]. Thus, one problem of the supply theory of money is

* Manuscript received January 12, 1960.

[1] This paper forms a part of a project financed by the Bureau of Business and Economic Research at UCLA. I also wish to thank the WDPC at UCLA for the use of their computing facilities.

[2] The schemata referred to in the text are J. E. Meade, "The Amount of Money and the Banking System," in *Readings in Monetary Theory* (New York: Blakiston, 1951), A. P. Collery, "A Graphical Analysis of the Theory of the Determination of the Money Supply," *Journal of Finance*, Vol. XI, 1956, pp. 328-32, and A. H. Meltzer, "The Behavior of the French Money Supply: 1938-1954," *Journal of Political Economy*, Vol. LXVII, 1959, pp. 275-96. Meade's and Collery's structures could easily be completed by the addition of semantic rules and specifications of the parameters. Such completions result in testable hypotheses. Another macro-schema was discussed in my paper, "A Case Study of U.S. Monetary Policy: Reserve Requirements and Inflationary Gold Flows in the Middle 30's," *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 94 Jahrgang, 1958, pp. 160-201. Semantic rules were stated when the schema was used to organize the data bearing on observable money stock patterns, but no behavior restrictions for the parameters were formulated. The schema was also used as a basis for the formulation of empirical money supply functions.

to complete the transformation of such schemata into useful scientific hypotheses. Associated with this completion is the evaluation of their predictive or explanatory power.

Monetary analysis contains fragments of analysis and isolated descriptive statements bearing on the micro-structure of the monetary system. A second problem is to weave this material relating to single banks and their interconnections into a coherent schema which provides a logical basis for macro-analysis.

This paper deals with the second problem. One of the procedures available will be applied in this paper to construct a schema which can be used to formulate an empirical macro-supply function of money within the institutional framework of the American banking system.[3]

A bank's supply decision with respect to new funds, measured by the desired rate of change of its asset-portfolio, is related to its volume of "available reserves," particularly the amount of surplus reserves. This relation is formulated with the aid of the "loss coefficient," a concept derived from a detailed consideration of the limitations imposed on a bank's portfolio adjustments to the existing volume of surplus reserves. The supply function of new bank funds describes the absorption of surplus reserves through the portfolio adjustments induced by the existence of such reserves. A supplementary relation summarizes the process generating surplus reserves. This second relation associates the surplus reserves with changes in the monetary base. Aggregation and combination of the two relations lead to a macro-schema which connects the money stock with policy variables. This macro-schema guides the specification of an empirical money supply function. Such a supply hypothesis of the money stock is prerequisite to any appraisal of monetary policy.

2. MICRO-THEORY: THE SUPPLY OF NEW BANK FUNDS

The construction of a bank's supply function of funds begins with the definition of "available reserves," v, and "surplus reserves," s. The variable v is the difference between a bank's total cash assets

[3] Another aggregation procedure based on matrix multiplier analysis is constructed in the author's forthcoming paper, "The Micro-Structure of the Monetary System and the Derivation of an Aggregative Money Supply Function." A second paper, "Income, Money Stock and the Loan Market," bases the derivation of the money supply function on a set of behavior functions describing the markets linking the monetary system with the income generating process. The analysis of both papers is based on the concepts developed in Part 2 of the present paper. It should be emphasized that the construction of the schema is guided by the existing institutions of the U.S. monetary system, but the resulting formalization can be easily adjusted to different institutional arrangements.

and its required reserves. The magnitude s is defined in terms of v and a (stock) demand function for available cash assets. Let v^d designate a bank's desired amount of reserves or cash assets, the amount the bank is willing to hold, given a set of specified conditions. We postulate the following stable cash asset demand function for banks:

$$(1) \qquad \begin{aligned} v^d &= w_0 + w_1 d_1 + w_2 d_2 , \\ w_0 &= w_0(r, b, a) , \end{aligned}$$

where d_1 are net demand deposits, d_2 are time deposits, w_1 and w_2 are the marginal propensities to hold cash assets with respect to the two types of deposits, r is an index function of loan rates and bond yields, b is the amount borrowed by the bank from Federal Reserve Banks, and a is the loan-investment ratio in the bank's portfolio. The derivative of w_0 with respect to r is always strictly negative and finite; the derivative is positive with respect to b and a. The postulate also includes the inequalities $w_1 > w_2 > 0$ and $v^d > 0$ for all positive values of r. The magnitude "surplus reserves" can now be introduced by an explicit definition:

$$(2) \qquad s = v - v^d .$$

The surplus reserves measure the volume of cash assets a bank desires to eliminate through asset expansion or to acquire via asset contraction. The banks adjust their portfolios continuously in such a manner that any s emerging will be absorbed. The connection between these portfolio adjustments and s forms the subject of the supply theory of bank funds, formalized by a suitably constructed supply function. This function is specified in terms of the concept of a loss coefficient.

In order to obtain an explicit definition of the loss coefficient, λ, we have to consider in some detail what happens to a bank's surplus reserves as a result of any given asset-expansion, x. A classification of the possible forms in which surplus reserves can be drained away whenever earning assets are increased provides a useful basis for the construction of the defining formula:

i. A portion of the deposits created by x will be withdrawn in the form of currency by the borrowers or the payees receiving the borrowers' checks.

ii. A portion of the deposits created by x will be redeposited at other banks. Not all payees receiving the borrowers' checks will deposit at the expanding bank.

iii. A portion of the deposits created by x may spill over into time

deposits and thus raise required and desired reserves against time deposits.

iv. Some of the deposits created by x may remain with the expanding bank and thus force an increase in required and desired reserves against demand deposits.

We shall now formalize these loss components. The following symbols are presented for this purpose:

c = spill-over into currency per dollar of x,
t = spill-over into time deposits per dollar of x,
p_1 = proportion of x maintained by borrowers on their deposit accounts due to minimum balance requirements of the bank,
p_2 = checks drawn against x and redeposited at the same bank per dollar of x,
p_3 = checks drawn against x and redeposited at other banks per dollar of x,
r^d = reserve requirements against net demand deposits,
r^t = reserve requirements against time deposits.

The first five magnitudes satisfy the following equation:

$$c + t + p_1 + p_2 + p_3 = 1 ,$$

implied by the definitions of the magnitudes involved.

The reserve requirements are always known to any bank with certainty. The bank is also assumed to know with certainty its own mind, at least in so far as w_1, w_2, and p_1 are concerned. With respect to the other magnitudes, the bank is stipulated to operate under incomplete information of the stochastic type. Long series of past observations enable the bank, in principle at least, to estimate the probability distribution of c, t, p_2, and p_3. In general, a bank will also have some hypotheses bearing on the mean of these distributions[4]. All stochastic variables are supposed to be replaced by their means in the subsequent constructions.

The elements introduced so far permit an immediate formalization of the first and third loss components:

cx is the currency outflow generated by x,
$(r^t + w_2)tx$ is the loss in surplus reserves due to an increase in time deposits associated with x.

[4] The mean of p_3 will probably be inversely related with the size of the bank and the number of its branches. Unfortunately, I find only the most casual evidence publicly available to support this hypothesis. It seems, for instance, that p_3 for the Bank of America has a mean of about .5. A medium sized bank in Southern California has a mean p_3 equal to .85, and a few small banks have a mean p_3 of approximately .93. This information does not support the "rule of size" formulated by Angell-Ficek and Shaw.

The specification of the other two components requires some additional considerations. The loss in surplus reserves associated with the redepositing of created deposits at other banks will, in general, not be equal to $p_s x$. The collection of checks regularly involves crediting and debiting of interbank deposits. The extent to which various channels are used for the collection of checks drawn against the expanding bank has a significant influence on the loss of surplus reserves resulting from any given asset expansion. In order to incorporate these institutional features into our scheme, three additional magnitudes are introduced:

f_1 = proportion of checks redeposited at other banks, drawn against deposits created at the expanding bank, whose collection involves a transfer of Federal Reserve funds,

f_2 = proportion of checks redeposited at other banks, drawn against deposits created at the expanding bank, collected by debiting the expanding bank's balances at other banks,

f_3 = proportion of checks redeposited at other banks, drawn against deposits created at the expanding bank, collected by crediting the other banks' deposit accounts at the expanding bank.

It is immediately obvious that these definitions imply

$$f_1 + f_2 + f_3 = 1 .$$

The information-postulate covering c, t, and p_i also applies to the f-factors.

When checks are drawn against x and redeposited at other banks, no surplus reserves will be lost—via this redepositing action—if all the checks are collected by crediting the other bank's deposits at the expanding bank, i.e., $f_3 = 1$. Only if $f_1 + f_2 > 0$, will such redepositing involve a loss in surplus reserves. Thus, we can now specify the second loss component with the following expression:

$$(1 - f_3)p_s x .$$

The fourth component depends on net demand deposits and not on total demand deposits. Consequently, variations in balances at other banks associated with the asset expansion via the f-factors must be explicitly accounted for. The remaining increase in demand deposits associated with x is given by

$$x - cx - tx - (1 - f_3)p_s x .$$

The change in the expanding bank's balances at other banks $f_2 p_s x$ must be added to the remaining increase in demand deposits. As a

result, the remaining increase in net demand deposits becomes

$$(1 - c - t - f_1 p_3)x .$$

By multiplying this expression into $(r^d + w_1)$, we obtain the increase in required and desired reserves against demand deposits associated with the asset expansion x. The total loss, l, in surplus reserves, originated by asset expansion, x, is the sum of the four components. So we have

$$(3) \quad l = [c + (r^t + w_2)t + (1 - f_3)p_3 + (r^d + w_1)(1 - c - t - f_1 p_3)]x .$$

The loss coefficient, λ, measuring the loss in surplus reserves per dollar of asset expansion, is then defined by the bracketed expression in (3).[5]

All the concepts required for the construction of the supply function of bank funds have now been introduced. Let a^s designate the supply quantity of such funds; then we define a^s by the following relation:

$$(4) \qquad a^s = \frac{1}{\lambda}(v - v^d) = \alpha(v, r, b, a, d_1, d_2, \lambda) ,$$

i.e., surplus reserves multiplied by the bank's expansion coefficient (=reciprocal of the loss coefficient) determine the current rate of change in the bank's asset portfolio and thus the supply of funds. The sign of the surplus reserves determines the sign of a^s. According to this construction, the properties of the bank's fund supply function reflect the properties of the bank's stock demand function for cash assets. The derivatives of α follow immediately from the derivatives of v^d. Two of them are of basic importance: $\delta\alpha/\delta v$ and $\delta\alpha/\delta r$. The effectiveness of monetary policy hinges crucially on these two properties. The reaction patterns of α with respect to b and a formalize standard assertions of Federal Reserve publications and the money and banking literature. These assertions are plausible enough, but so far there has been no systematic evaluation of their validity. The money supply schema to be constructed provides a framework which will enable us to appraise the relative significance of variations of bank indebtedness "b" and the loan-investment ratio "a" in the money supply mechanism and to assay also the v^d-function's interest elasticity[6].

[5] Appendix I investigates the properties of λ in some detail.

[6] It has been asserted in the literature, particularly by Shaw in *Money, Income and Monetary Policy* (Chicago: Irwin, 1950), p. 132, that banks may be forced by inadequate and interest-inelastic loan demand to hold undesired surplus reserves. This thesis thus claims that under specifiable conditions banks are "pushed off" their cash asset demand function. The postulated properties of the v^d function imply that some arrangement to prevent banks from competing for assets by appropriate adjustments in pertinent prices is a necessary condition for the persistence of surplus reserves. The existence of a

A supplementary relation is constructed in this section which explains the conditions generating surplus reserves for a given bank. This supplementary relation forms an essential building block in the construction of a useful money supply function. All the events and magnitudes which shape the emergence of s are classified in eight categories. These will be discussed presently, and a formal expression will be associated with each category. The sum of these expressions will provide an exhaustive statement of the processes generating surplus reserves.

i. Currency flows in and out of banks for reasons other than changes in the asset portfolio and associated changes in deposits. Other factors, besides deposits, determine the public's demand for currency[7]. Thus, we isolate a currency flow independent of changes in portfolio and deposits. We denote this net currency inflow into the bank by n_1. The volume of s generated by such a net inflow is given by the expression

$$[1 - (r^d + w_1)(1 - g_2)]n_1 ,$$

where g_2 is the marginal adjustment in the bank's balances at other banks to the change in deposit liabilities associated with n_1 flows.

ii. The formulation of the loss coefficient assumed the existence of a demand function for time deposits on the part of the public. This demand function contains other variables besides demand deposits as arguments. Variations in these other magnitudes generate conversions

market for government securities in which banks operate competitively with other economic units renders the thesis rather dubious. Economic theory assigns a negative interest elasticity to the public's stock demand for such securities and we have no basis in *monetary* theory to deny this. Consequently, banks are apparently always able to unload (or acquire) surplus reserves through operations on the bond market. Still, the thesis requires a careful formalization and evaluation. The macro-schema eventually to be constructed provides us with an instrument for such appraisal.

[7] We postulate the existence of a stable demand function for currency; $C^p = c_0(z) + c_1 M$. c_1 is a function of the banks' currency spillover rates contained in their loss coefficient. The nature of this function defines the aggregation procedure involved. z is a vector of variables, and c_0 a function with z as argument. An investigation of some aspects of this first component of the public's demand function for currency was presented by S. L. MacDonald, "Some Factors Affecting the Increased Relative Use of Currency since 1939," *Journal of Finance*, Vol. XI, 1956, pp. 313-27. Consider also Phillip Cagan's "The Demand for Currency Relative to the Total Money Supply," *Journal of Political Economy*, Vol. LXVI, 1958, pp. 303-28. Cagan's demand function differs in several respects from the one used here. But his study indicates the most important magnitudes to be included under z. The currency flows designated by n_1 are related to variations in c_0. A given change, Δc_0, is equal to the sum of the n_1 taken over all banks. Thus, an individual bank's n_1 depends on Δc_0 and its "relative position" in the system.

of demand into time deposits or vice versa, i.e., we admit the existence of a flow of funds between the two classes of deposits independent of changes in the asset portfolio and associated deposit changes[8]. Let n_2 denote the net conversion of demand into time deposits at the given bank; then the volume of surplus reserves emerging is expressed by

$$[(r^d + w_1)(1 - g_2) - (r^t + w_2)]n_2 .$$

iii. Checks are drawn against already existing deposits of a bank and redeposited elsewhere. These checks are currently balanced by checks drawn on other banks and deposited at the given bank. The difference between the accrual of checks drawn against other banks and checks drawn against existing deposits at the given bank and redeposited at other banks is denoted by n_3. The n_3 flows in the system are actually generated by two distinct mechanisms: One is the redistribution of already existing deposits over the system of banks; the other is the redistribution of new funds issued by the Federal Reserve over the system. Such a redistribution of new reserves accompanies every change in some asset portfolio. Consequently, a positive n_3 of some bank may result either from a shift in existing deposits to this bank or from asset expansions at other banks. In general, the larger the asset expansion at other banks, the greater will n_3 tend to be. The volume of s obtained from n_3 is given by

$$[(1 - g_3) - (r^d + w_1)g_1]n_3 ,$$

where g_1 and g_3 specify the proportion of the "clearing balance" collected via the Federal Reserve mechanism and via debiting the receiving bank's liabilities to other banks.

iv. Surplus reserves will emerge from transactions between the Federal Reserve Banks and the bank under consideration. Two types of transactions must be distinguished.

n_4 is the net accrual of cash assets associated with transactions between the bank and Federal Reserve Banks which involve either the bank's other (i.e., non deposit) liabilities (e.g., borrowing from the Federal Reserve) or the bank's earning assets.

[8] Similar to the currency case, the demand function for the time-deposits is postulated to be

$$T = t_0(y) + t_1 M .$$

t_1 is an aggregate of the bank's spillover rates into time deposits. y is again a vector; it may contain interest rates on time-deposits, other interest rates, service charges on checking deposits, real wealth, etc. The sum of n_2 flows in the system is equal to the variations in t_0, according to the formula stated in Section 3a of the text.

n_3 is the net accrual of cash assets associated with transactions between the bank and the Federal Reserve which involve the bank's deposit liabilities with the exception of other banks' accounts at the given bank. Examples: Transfer of Treasury funds from tax and loan account to Federal Reserve account; open-market operations which involve buying or selling by the public.

With these specifications, the contribution of this class of transactions to s is given by the expression

$$n_4 + (1 - \nu)n_3 ,$$

where ν depends on the assumption to be made concerning the behavior of the deposits generated not by changes in the asset portfolio but by transactions with the Federal Reserve Banks. If we assume that the same pattern applies to both types of deposits, we have $\nu = \lambda$. The minimum assumption would be that $\nu = (r^d + w_1)(1 - g_2)$.

v. Changes in reserve requirements form a major source of s. The volume of surplus reserves which emerges under this heading is measured by

$$-\Delta r^d d_1 - \Delta r^t d_2 .$$

vi. A bank's total demand for cash assets was specified by a demand function. Within a given total, variations may occur among the two major components. Such allocative variations between Federal Reserve money and balances at other banks occur partly in the form of marginal reactions to a bank's deposit changes. Previous specifications determine this portion of a bank's cash asset pattern[9]. Another portion reflects reallocations within the demand component w_0. The existing definition of "net demand deposits" implies that these changes in the composition of w_0 contribute to the current rate of s generation. The effect on s of a bank's reallocation of cash assets from Federal Reserve money to banker's balances which occurs independently of deposit changes—measured by Δh_0—is stated by the expression

$$(r^d + w_1)\Delta h_0 .$$

vii. A bank may be the recipient of Federal Reserve funds resulting from other banks' reallocation of their components of w_0. Let d_0^b indicate the bank's deposits which are balances included in other banks w_0. Actually d_0^b is a component of d_1. The accrual of s resulting in

[9] The previous specifications determine the following change in a bank's balances at other banks associated with changes in its deposit liabilities:

$$g_2(n_1 + n_2 + n_3) - f_2 p_3 k s .$$

this manner is given by $[1 - (r^d + w_1)(1 - g_2)]\Delta d_0^a$.

viii. Absorption and release of s associated with variations in v^d resulting from a^s operate via the marginal propensities w_1, w_2 and have been included in the loss coefficient. Variations in w_0 occur independently of a^s and concomitant deposit changes. These variations immediately generate s. Thus, the last category contributing to surplus reserves is $-\Delta w_0$.

The above eight components completely determine s by means of the following relation:

$$
\begin{aligned}
&[1 - (r^d + w_1)(1 - g_2)]n_1 + [(r^d + w_1)(1 - g_2) - (r^i + w_2)]n_2 \\
&\quad + [(1 - g_2) - (r^d + w_1)g_1]n_3 + n_4 + (1 - \nu)n_5 \\
&\quad - \Delta r^d d_1 - \Delta r^i d_2 + (r^d + w_1)\Delta h_0 \\
&\quad + [1 - (r^d + w_1)(1 - g_2)]\Delta d_0^b - \Delta w_0 = s\,.
\end{aligned} \tag{5}
$$

Relations (4) and (5) together constitute our basic framework. Relation (5) describes the emergence of s, and relation (4) describes the bank's reaction to s in the form of a demand for earning assets designed to absorb this s.

3. MACRO-THEORY: CONSTRUCTION OF THE AGGREGATIVE SCHEMA

The discussion of the single bank's position may be used to build a theory which contains some general propositions concerning observable patterns of individual banks. This is not a major purpose in monetary theory. Relations (4) and (5) are important because they provide a framework guiding the construction of an aggregative schema which will be useful for the specification of an empirical money supply function.

Traditionally, the system's reaction pattern is derived from the single bank's formula either by means of a reinterpretation of the component magnitudes in the loss coefficient or with the aid of a rudimentary dynamic analysis. This reinterpretation proceeds in the following way: The various components are more or less explicitly redefined relative to the system, and the system is treated as if it were a single bank. As a result of this reinterpretation, p_3 practically vanishes. Consequently, in this manner, the system's loss coefficient would be

$$\lambda^s = c + (r^i + w_2)t + (r^d + w_1)(1 - c - t)\,. \tag{6}$$

The reinterpretation procedure obtains, in general, a λ^s with an obscure and tenuous meaning in terms of the system's micro-structure.

Analytically more satisfactory aggregation processes are at our disposal. One is based on a matrix multiplier analysis, and the other

considers a suggestion made by Philipps but first formalized by Rogers.[10] The latter approach, explored in this paper, constructs the aggregate expansion coefficient as the limit of a series whose terms describe the single bank's adjustments to its surplus reserves. The derivation of the aggregative formula proceeds in the following manner: Let the system's surplus reserves at a given point in time be s_0. It is assumed that all surplus reserves are always located with one bank and all the banks' expansion coefficients are identical. The bank possessing s_0 will expand assets by an amount ks_0. The volume of checks drawn against the first bank and redeposited at a second bank will be p_2ks_0. The public's deposits at the second bank change by the same amount. This magnitude will differ in general from the accrual of cash assets to the second bank, which is equal to the increase in total deposits expressed by $(1 - f_2)p_2ks_0$. All the f-factors in the formulas have the established meaning, i.e., they are defined relative to the expanding bank. We consider now the reserve drains occurring at the second bank associated with the deposit gains resulting from the first bank's asset expansion. This formalization depends on our choice of a basis for the spillover rates c_2 and t_2. They may be related to total deposit accrual $(1 - f_2)p_2ks_0$ or to the public's deposit gains p_2ks_0. We choose the latter basis. The sum of currency drain and drain into required and desired reserves against demand and time deposits is expressed by

$$c_2p_2ks_0 + (r^t + w_2)p_2ks_0 + (r^d + w_1)(f_1 - c_2 - t_2)p_2ks_0 .$$

These reserve drains must be subtracted from the cash assets received by the second bank. The difference is the volume of surplus reserves in the system after the first round of asset expansions:

$$(7) \qquad s_1 = [1 - f_2 - c_2 - t_2(r^t + w_2) - (r^d + w_1)(f_1 - c_2 - t_2)]p_2ks_0 .$$

Given any s_0, the formula determines the unfolding sequence of surplus reserves. Associated with this sequence is one specifying the changes in the money stock. In order to express these changes, we include time deposits in the money stock, but exclude interbank deposits. The

[10] J. H. Rogers, "The Absorption of Bank Credit," *Econometrica*, Vol. 1, 1933, pp. 63–70. Rogers was the first to obtain the system's expansion coefficient k explicitly as the limit of a geometric series describing successive reactions of single banks to the continuous redistribution of surplus reserves. But this derivation is made under highly restrictive assumptions which rule out all currency flows. Shaw presents on p. 140, *op. cit.*, a short verbal discussion of the expansion process in a two-bank system. At every step of the process, one of the banks expands according to its k and the surplus reserves obtained via the clearing. Shaw seems to think that this process necessarily leads to a result expressed by the k^s obtained by reinterpretation of the single bank's k. Comparison of $(\lambda - \mu)$ with λ^s reveals that this is in general not the case.

addition to the money stock which results from ks_0 appears at both the expanding bank and the second bank which receives the redeposited checks. The addition made at the first bank is measured by $(1 - p_2)ks_0$, and the addition made at the second bank by p_2ks_0. Their sum is ks_0, i.e., equal to the asset expansion. Consequently, the increase in the money stock associated with any s_t is $\Delta m_t = ks_t$. The total amount of money created by s_0 is then the sum of these additions made in each round of asset expansion. Consequently, the total addition is expressed by

$$\Delta m = \sum_{t=0}^{\infty} \Delta m_t = k\left[1 + \sum_{t=1}^{\infty} (\mu k)^t\right] s_0 , \tag{8}$$

where

$$\mu = [1 - f_2 - c_2 - t_2(r^i + w_2) - (r^d + w_2)(f_1 - c_2 - t_2)]p_2 . \tag{9}$$

Two postulates are introduced at this point, and they both form essential steps in the process leading to the money supply function. First, μk is stipulated to be a positive number in the open interval $(0, 1)$. This restriction assures a non-oscillatory convergence of the series. Secondly, it is stipulated that the system's adjustment process to any s is sufficiently rapid relative to the time interval for which our data are defined. The speedy absorption of surplus reserves by the system enables us to approximate observable reactions with the aid of the limit of the series.

An aggregative relation which corresponds to the micro-relation (4) immediately emerges from these two postulates. This aggregative relation specifies the *total* rate of change of the money stock induced by the system's surplus reserve position σ. Thus, we formulate the first component of an aggregative schema describing the system's behavior[11].

[11] The implausibility of some of the assumptions used to derive (9) may detract attention from the empirical usefulness of the supply schema ultimately constructed. The prevalence of the "fallacy of denying the antecedent" and the misconceived idea of uniqueness with respect to sufficient conditions in the literature necessitate two explanations:

(a) A valid deduction transfers the truth of the premises to the conclusion, but it does not transfer falsehood.

(b) The assumptions were chosen for their simplicity. The linear succession of single banks, the concentration at any moment of time of the system's surplus reserves at one bank and the equality of all banks' pertinent coefficients can be dismissed from our analysis. Even with the admission of a simultaneous movement of bank expansions, of a wide distribution of surplus reserves over the system and significant differences between individual banks' behavior characteristics, we can still derive the multiplier formula $k^1/(1 - \mu k^2)$. k^1 and k^2 are in general distinct averages of all banks' expansion coefficients and μ an average of all the banks' μ_i. This result will be established in a forthcoming paper mentioned in footnote 3.

$$\dot{M} = \frac{k}{1 - \mu k}\sigma \,, \quad \text{or} \quad \dot{M} = \frac{1}{\lambda - \mu}\sigma \,. \tag{10}$$

A discussion of some basic properties of the monetary multiplier $(\lambda - \mu)^{-1}$ is contained in Appendix II.

We now formulate an aggregative counterpart to the second micro-relation given by equation (5). Aggregation proceeds by simple summation after a suitable adjustment of these relations. The system's total volume of surplus reserves σ is determined by the expression,

$$\sigma = -a_1\dot{c}_0 + a_2\dot{t}_0 + l^1 + a_3\dot{B} + l^2 + a_4\dot{\varepsilon}_0 + \dot{\gamma}_0 - \dot{w}_0 \,. \tag{11}$$

The following discussion describes the connection between (5) and (11):

a. The first two components describe how currency flows and conversions into time deposits independent of portfolio changes contribute to s. The variables c_0 and t_0 have already been introduced as the first components in the public's demand functions for currency and time deposits. The coefficients a_1 and a_2 are specified as averages of the pertinent components in (5). The defining relations are

$$a_1 = -\sum_{i=1}^{n} \frac{1 - (r^{di} + w_1^i)(1 - g_{21})}{\dot{c}_0} n_1^i \,;$$

$$a_2 = \sum_{i=1}^{n} \frac{[(r^{di} + w_1^i)(1 - g_{21}) - (r^{ti} + w_2^i)]}{\dot{t}_0} n_2^i \,;$$

where

$$-\dot{c}_0 = \sum_{i=1}^{n} n_1^i \,; \qquad \dot{t}_0 = \sum_{i=1}^{n} n_2^i \,.$$

b. The third component of the micro-relations has already been partly aggregated by the multiplier equation (10). It has been previously remarked that the flow components n_3 consist of two distinct subcomponents. The subcomponent resulting from the redistribution of cash assets associated with individual banks' asset expansion forms a part of the multiplier mechanism summarized by $(\lambda - \mu)^{-1}$. The other subcomponent, resulting from the redistribution of already existing deposits among different banks still needs aggregation. We designate with n_3^{ii} the subcomponent under consideration for the ith bank. The sum of all n_3^{ii} is by construction zero. The magnitude l^1 introduced by the definition

$$l^1 = \sum_{i=1}^{n} [(1 - g_{31}) - (r^{di} + w_1^i)g_{11}]n_3^{ii}$$

measures the rate at which surplus reserves are generated over the system by the current redistribution of already existing deposits.

The contribution of this mechanism to σ depends on the existence of significant differences in the bracketed expressions for different banks. Such differences are not a sufficient, but are a necessary, condition for this mechanism to affect σ. Also, these differences are not uniquely determined by differences in reserve requirements. A systematic contribution to σ can be expected if there is a positive association between the signs of the n_3^{ii} flows and the relative size of the bracketed expressions. In particular, if banks with positive n_3^{ii} tend to possess larger magnitudes $[(1 - f_{ii}) - (r^{di} + w_i^j)f_{ii}]$ than banks with net outflows, then this deposit redistribution mechanism generates surplus reserves over the system[12].

c. Summation of the fourth component results in the expression

$$\sum_{i=1}^{n} [n_4^i + n_3^i(1 - \nu^i)] .$$

The sum of n_4^i and n_3^i is, by definition, equal to the rate of change of the monetary base B. The macro-component $a_3\dot{B}$ is equal to the above sum of micro-components provided the following equations are satisfied:

$$a_3 = v + (1 - v)(1 - \nu) ,$$

$$v = \frac{\sum n_4^i}{\sum (n_4^i + n_3^i)} , \qquad 1 - \nu = \frac{\sum (1 - \nu^i) n_3^i}{\sum n_3^i} ,$$

where v measures the portion of $\dot{B}$ which accrues to banks without simultaneously raising their deposits.

d. The fifth component, l^3, describes the volume of reserves currently liberated (or frozen) by changes in reserve requirements. This magnitude is the sum of all corresponding components of the micro-relations according to the definition

$$l^3 = - \sum (\dot{r}^{d1} d_1 + \dot{r}^{t1} d_2^i) .$$

The sum of l^1 and l^2 is denoted by l^3: This quantity measures the current rate of σ generated via fiat changes in reserve requirements or changes in the system's average reserve requirements induced by shifts in the relative position of single banks associated with the deposit

[12] The literature on money and banking repeatedly asserts the possibility of serious volatile shocks exerted on the money stock by this deposit redistribution process. This possibility certainly exists. Contrary to a prevalent psychological reasoning in economics, accepting such a possibility has no implication as to its relevance. Computations of the effect of l^1 on M have shown two important features: i) The effect is very small, e.g., the annual rate of change in M attributable to this process is at most 0.2 percent in the last fifteen years and as a rule less than 0.1 percent; ii) The operation of this process on M exhibits a strongly marked pattern of regularity.

redistribution mechanism.

e. The sixth term of the macro-relation expresses the contribution of variations in the system's interbank deposit structure to the generation of surplus reserves. The coefficient a_4 is the sum $\sum_{i=1}^{4}(r^{di} + w_1^i)$ of the average reserve requirements against demand deposits and average marginal propensities of the Federal Reserve System's four major groups of commercial banks. The magnitude $\dot{\varepsilon}_0$ denotes an average of the rate of change in the four groups of banks' net position on interbank deposit account[13].

The last two terms are sums of the corresponding individual components. Thus, $(\dot{w}_0 - \dot{\gamma}_0)$ designates the rate of change of the banks' demand component for Federal Reserve money in excess of required reserves, which is independent of their deposit liabilities and dependent on interest rates.

The multiplier relation (10) only determines the rate of change in M associated with the bank's portfolio adjustments to a given surplus reserve position. The discussion of the components in the supplementary relation indicates that there exist transactions, denoted by n_s, which

[13] All banks are divided into four categories: Central Reserve city banks (1), Reserve city banks (2), country banks (3), and non-member banks (4). For each group of member banks we observe a unique reserve requirement r^{di}, $i = 1, 2, 3$. The reserve requirement associated with the non-member banks must be specified as an average

$$r^{d4} = \frac{\sum_{j=1}^{n_4} r^{dj}(h_0^j - d_0^{bj})}{\sum_{j=1}^{n_4} (h_0^j - d_0^{bj})},$$

with the sums taken over all n_4 non-member banks. An average marginal propensity w_{1s}, $s = 1, 2, 3, 4$, is defined analogously for each group. Each group's position on interbank deposit account is defined by

$$h_0^s - (1 - g_{1\cdot}^s)d_0^{bs} = \sum_{j=1}^{n_s} [h_0^{sj} - (1 - g_{1\cdot}^{sj})d_0^{bsj}],$$

the summation running over all banks j in group s. Thus we ultimately specify

$$a_4 = \sum_{s=1}^{4} (r^{ds} + w_1^s), \quad \text{and} \quad \varepsilon_0 = \frac{\sum_{s=1}^{4} (r^{ds} + w_1^s)[h_0^s - (1 - g_{1s})d_0^{bs}]}{\sum_{s=1}^{4} (r^{ds} + w_1^s)}$$

The macro-variable γ_0 is defined by the sum $\sum_{i=1}^{n} d_0^{bi}$. With these specifications, it is easy to establish that

$$a_4\dot{\varepsilon}_0 + \dot{\gamma}_0 = \sum_{i=1}^{n} (r^{di} + w_1^i)\Delta h_0^i + \sum [1 - (r^{di} + w_1^i)(1 - g_{1i})] - \Delta d_0^{bi},$$

i.e. the sum of the two macro-components is equal to the sum of the corresponding micro-components taken over all banks. The aggregation procedure thus separates the volume and the distribution patterns of interbank deposits into distinct terms of the macro-relation.

generate changes in the money stock independent of portfolio adjustments.

The current rate of this direct contribution to the money stock is measured by $(1 - v)\dot{B}$. This magnitude is added to the original multiplier relation. We obtain

$$\dot{M} = \frac{1}{\lambda - \mu}\sigma + (1 - v)\dot{B} . \tag{12}$$

Equation (13) follows immediately from (12) by replacing σ with the aid of (11)

$$\dot{M} = -m_1\dot{c}_0 + m_2 t_0 + m_0 l^3 + m_3\dot{B} + m_4\dot{\varepsilon}_0 + m_0(\dot{\gamma}_0 - \dot{w}_0) , \tag{13}$$

$$m_0 = \frac{1}{\lambda - \mu} , \quad m_i = m_0 a_i , \quad i = 1, 2, 4 .$$

The coefficient m_3 is defined by

$$m_3 = m_0[v + (1 - v)(1 - \nu)] + (1 - v) = m_0 + (1 - v)(1 - \nu m_0) .$$

The second component in m_3 vanishes as v tends to one. Thus, for $v = 1$, m_3 coincides with m_0. If both ν and v are zero, then $m_3 = m_0 + 1$; if $\nu = 1$ and $v = 0$, then $m_3 = 1$. This last case would not be consistent with the available evidence on m_3. Evidence also suggests that $0 < v < 1$. It is assumed that $\nu = m_0^{-1}$. This condition eliminates the second component, so that $m_3 = m_0$. The resulting simplification permits the combination of $m_0 l^3 + m_3\dot{B}$ into a single expression $m_0(\dot{B} + l^3)$.

Equation (13) states a formal construct for the discussion of the growth rate of the money stock. The usefulness of such schemata depends on the degree of corroboration of the hypotheses to be erected on their basis. Such hypotheses require definite restrictions on the coefficients m_i. Previous assumptions combined with the process of constructing m_0 and m_4 establish the existence of definite values for these coefficients. This is not the case for a_1 and a_2 (and thus for m_1 and m_2). It is therefore postulated that the flow patterns underlying these coefficients exhibit a persistent structure, which would assure approximate constancy for a_1 and a_2 for all $\dot{c}_0 \neq 0 \neq \dot{t}_0$. In case $\dot{c}_0 = 0 = \dot{t}_0$ the coefficients are indeterminate by construction. This indeterminacy is removed by arbitrarily assigning the values of a_1 and a_2 obtained for $\dot{c}_0 \neq 0 \neq \dot{t}_0$ to the case of $\dot{c}_0 = 0 = \dot{t}_0$. This removal of a possible indeterminacy of m_1 and m_2 is consistent with the aggregation procedure.

A schema for the specification of a money supply function can now

be obtained by term-wise integration of the growth rate expression. However, this process raises some difficulties. All the coefficients m_j $(j = 1, \cdots, 4)$ are functions of the reserve requirements. As it happens, reserve requirements are constant over lengthy intervals. This enables us to integrate the growth functions at least over these intervals. The result is:

$$M = a - m_1 c_0 + m_2 t_0 + m_0(B + L^s) + m_4 \varepsilon_0 + m_3(\gamma_0 - w_0) , \tag{14}$$

where a is the integration constant.

A reduction in r^d and r^t raises m_0 and m_1 and very likely contracts m_2 and m_4. The multiplier's sensitivity to variations in r^d has been evaluated in Appendix II. It is shown that a .01 change in r^d means approximately a 2.5 percent change in m_0. Thus, with the exception perhaps of 1936–37, differences in m_0 associated with variations in r^d can, in general, not be discerned from sampling fluctuations in the estimates of m_0. Such obliteration of the systematic connection between m_0 and r^d suggests an extended application of equation (14) to cover intervals containing fiat changes in reserve requirements, provided these changes do not exceed .03 in numerical value.

The money supply function explains the observable behavior of the money stock in terms of a specified combination of magnitudes. Such a function, expressing a confirmable hypothesis, is obtained from the schema (14) with the assignment of semantic rules to the variables and definite values to the coefficients. The selection of these rules and the derivation and testing of coefficient values will not be discussed in this paper. Still, the schema constructed is sufficiently specified to describe some general features of the money supply mechanism.

The multiplier m_0 dominates all the coefficients in the schema. It describes the reaction in M to a unit change in the monetary base. The same reaction is predicted for a unit change in the total volume of interbank deposits or for every unit of reserves liberated by changes in reserve requirements. The definition of m_0 established that $m_0 > 1$. The construction of the schema also implies the following order of magnitude relations among the coefficients:

$$m_0 > m_1 > m_2 , \quad \text{and} \quad m_0 > m_4 > m_2 .$$

These inequalities follow immediately from the definitions of the a_i which multiply m_0 to obtain m_i, $i = 1, 2, 4$. The schema is insufficiently specified to determine the comparative order of m_1 and m_4. The two coefficients may very likely be similar.

Interbank deposits are a major institution inadequately integrated into traditional money supply theory. It is explicitly incorporated into

the schema in a manner which separates a structural and a scale effect on M of variations in interbank deposits. The structural effect is generated by changes in the distributional patterns of a constant volume γ_0 of interbank deposits. This volume is equal to the sum $\sum_{i=1}^{4} h_0^i$ of all balances at other banks of the three groups of member banks and the non-member banks. It is also equal to the sum $\sum_{i=1}^{4} d_0^{bi}$ of all interbank deposit liabilities of the four groups of banks. We have thus to distinguish two distinct distributional patterns, the h_0 and the d_0^b pattern. The effect on M exerted by any given change in each pattern is formulated by the two statements

$$dM_h = m_0 \sum_{i=2}^{4} [(r^{di} - r^{d1}) + (w_1^i - w_1^1)] \Delta h_0^i ,$$

$$dM_d = m_0 \sum_{i=2}^{4} [(1 - g_{2i})(r^{d1} + w_1^1 - r^{di} - w_1^i)] \Delta d_0^{bi} .$$

We notice that parallel changes in the h_0 and d_0^b distributions have opposite effects on M. An increase in balances of Central Reserve city banks compensated by a reduction in balances of Reserve city banks (i.e., $\Delta h^1 = -\Delta h^2 > 0$ and $\Delta h_3 = \Delta h_4 = 0$) raises the money stock. A reduction in other banks' deposits at Reserve city banks compensated by an increase in other banks' deposits at Central Reserve city banks (i.e., $\Delta d_3^1 = -\Delta d_3^2 > 0$ and $\Delta d_3^3 = \Delta d_3^4 = 0$) lowers the money stock. The formulas indicate that the structural effect depends crucially on the difference between reserve requirements and marginal propensities to hold cash assets relative to deposits. If we disregard the marginal propensities w_{1i} and w_{2i} then the existing r^{di} would imply approximate values of .02, .07, and .08 for the bracketed expressions. Estimates of m_0 in Table 2c associated with the money supply measure M^1 (currency outside banks and demand deposits adjusted) and M^2 (M^1 plus time deposits of commercial banks) permit an evaluation of the order of magnitude of the structural effect. A reallocation of $3 billion of interbank deposit liabilities from Central Reserve city banks in equal shares to the other three groups would raise M^1 approximately $.4 billion and M^2 by approximately $.6 billion. A one percent variation in the money stock M^1 would require a reallocation of interbank deposit liabilities of about $8 billion from Central Reserve city banks to country banks and about $28 billion from Reserve city to country banks. Reallocations of such magnitudes are extremely unlikely to occur within the span of one or two years. The more likely distributional changes would have a correspondingly smaller impact on the money stock. So far, we have neglected to consider the marginal propensities w_{1i} and

w_{24}. Casual inspection of data suggests the existence of an inverse relation between r^{44} and w_{14}. If this relation were to be substantiated we would probably have to conclude that the product of m_0 with any of the bracketed expressions is of very small order. In such a case even extremely unlikely variations in the distributional patterns would generate variations in M barely discernible from random fluctuations. Purely redistributional effects of interbank deposits could, under these conditions, be neglected.

The scale effect of variations in interbank deposits isolates the impact of changes in the total volume from changes in the distribution pattern. In this case the patterns of h_0 and d_0^{ij} are held constant. The scale effect is measured by the differential

$$dM = m_0(a_4\varepsilon_0 + \gamma_0)dg \ ,$$

where dg indicates the change in a scale number. Let $dg = \$1 \text{ billion}/\gamma_0$, so that dg is a pure number. Then we may write

$$m_0\left[a_4\frac{\varepsilon_0 \cdot \$1 \text{ billion}}{\gamma_0} + \$1 \text{ billion}\right] \sim m_0 \cdot \$1 \text{ billion} \ .$$

Thus a unit change in the total volume of interbank deposits γ_0 with constant h_0 and d_0^{ij} proportions would cause about the same reaction in M as would a unit change in the monetary base.

The derivation of (14) also reveals that the definition of net demand deposits determines the coefficients of ε_0 and γ_0. A redefinition of "net demand deposits" which equates them with demand deposits would eliminate ε_0 from the schema and modify the coefficient of γ_0 to

$$m_0\frac{\sum(1 - r^{44} - w_i^4)d_0^{3i}}{\sum d_0^{3i}} < m_0 \ .$$

The redefinition would therefore completely eliminate any structural effects of interbank deposits and compress their scale effect by approximately one-fifth. If γ_0 were to exhibit irregular and considerable variations, such a redefinition would improve the degree of control over M exerted by the monetary authorities.

The schema indicates no reason to expect the ratios M/B or $\Delta M/\Delta B$ to be approximately constant. The public's behavior expressed by c_0 and t_0 and the banks' behavior expressed by γ_0 and w_0 contribute jointly with B or $B + L^3$ to shape the movement of M. Substantial variations in the ratios indicated are quite consistent with any hypothesis based on our schema, provided the variations can be traced to c_0, t_0, γ_0, w_0 or L^3. Irregular variations in these ratios present no evidence for the "inherent instability of monetary affairs" but reflect

only an inadequately constructed hypothesis. Tables 1c and 2c contain a summary of statistical estimates related to equation (14). Table 1c compares the correlation coefficients of a simple hypothesis, which acknowledges only B as a determinant of M, and of extended hypotheses, which also consider the effect of c_0, t_0, and L^3 on M, for four disjoint sample periods. Such a comparison of correlation-coefficients is based on their interpretation as measures of the comparative predictive performance of the hypotheses under consideration[14]. The sample estimates thus establish that the neglect of c_0, t_0, and L^3 results in a hypothesis with poor and unreliable predictive power. Table 2c corroborates this conclusion. The relative stabilization of the estimates of the monetary multiplier drawn from sample periods with radically different "economic climate" supplies additional evidence for the comparative relevance of a hypothesis which explicitly recognizes the operation of currency demand, time deposit demand, and the cumulative effect of changes in reserve requirements in the explanation of M.[15]

4. CONCLUSION

We started with the concept of a demand function for cash assets, developed the concept of the loss coefficient and, with the aid of a series of additional assumptions, derived an aggregate money supply function. The money stock is explained in terms of some component of the public's demand functions for currency and time deposits, the monetary base adjusted for the cumulated reserve liberations, the

[14] See in this context H. Theil, *Economic Forecasts and Policy* (Amsterdam: North-Holland Publishing Co., 1958), pp. 211-12. Theil justifies the choice of maximum (multiple) correlation as a criterion of selecting hypotheses. He shows that this criterion will "lead us" in the average "to the correct choice" of hypothesis in terms of the ultimately pertinent criterion, viz. predictive or explanatory performance.

[15] The coefficients of equation (14) were specified for a money stock measure which includes time deposits. In case we desire to use a more restrictive measure excluding time deposits the coefficients require suitable adjustment. The construction of (14) determines the nature of these changes:

(a) The monetary multiplier m_0 changes to $[1 - (t_1 + t_2 p_3)](\lambda - \mu)^{-1}$.

(b) The relative changes in the monetary multiplier differ from those presented in Appendix II in the following manner:

i. The relative change with respect to p_3 is less by $t_2[1 - (t_1 + t_2 p_3)]^{-1}$.

ii. The relative change with respect to $(t_1 + t_2 p_3)$ is less by $[1 - (t_1 + t_2 p_3)]^{-1}$.

iii. The relative change with respect to all other components is numerically smaller according to the proportionality factor $[1 - (t_1 + t_2 p_3)]$.

(c) Equation (12) is to be replaced by $\dot{M} = m_0 \sigma + (1 + v)\dot{B} - \dot{t}_0$. This modification implies that the coefficient of t_0 in the adjusted money supply function is now $(m_0 a_2 - 1)$.

(d) Substitution of a less inclusive for a more inclusive money stock measure thus lowers the value of the multiplier and changes the sign of the t_0 coefficient.

interbank deposit structure, and a specific component of the banks' demand for Federal Reserve money. Whether or not this schema will actually perform usefully in explaining or predicting patterns of M cannot be decided by *a priori* plausible arguments. Such arguments have at best a suggestive force with respect to the *construction* of hypotheses, but they contain no relevant information with which to *assay* these hypotheses by processes of confirmation and testing. An adequate completion of money supply theory, based on the above schema, requires the selection of semantic rules to be associated with c_0, t_0, L^3, w_0, and γ_0. Closely connected with this choice of rules is our need for more detailed information concerning the patterns dominating these magnitudes, so as to suggest the formalization of suitable hypotheses capable of explaining their behavior. Thus, the schema obtained exhibits a frame guiding further investigations in the field of empirical money supply theory.

University of California (Los Angeles), U.S.A.

APPENDIX I: PROPERTIES OF THE LOSS-COEFFICIENT

The derivatives of λ reveal some major features of the loss coefficient, particularly how variations in the component magnitudes among different types of banks shape the comparative level of their loss coefficient:

$$\frac{\delta\lambda}{\delta c} = 1 - r^d - w_1 ; \qquad \frac{\delta\lambda}{\delta r^d} = \frac{\delta\lambda}{\delta w_1} = 1 - c - t - f_1 p_3 ;$$

$$\frac{\delta\lambda}{\delta r^t} = \frac{\delta\lambda}{\delta w_2} = t ; \qquad \frac{\delta\lambda}{\delta t} = (r^t - r^d) + (w_2 - w_1) ;$$

$$\frac{\delta\lambda}{\delta p_3} = f_1(1 - r^d - w_1) + f_3 .$$

The connection between λ and the f-factors cannot be obtained by simply differentiating λ with respect to f_i. This procedure presupposes that the f_i are independent of each other and that we may consider two to be constant and just one to vary. Their sum, however, is unity. Consequently, we differentiate both λ and the equation expressing this restriction, and so

$$d\lambda = -p_3 df_3 - p_3(r^d + w_1)df_1 ,$$
$$df_1 + df_2 + df_3 = 0 .$$

Actually, differentiation with respect to c, t, and p_3 is also subject to

a restriction, namely, $c + t + \sum p_i = 1$. We can safely assume that $p_1 + p_2 > 0$ and thus presuppose that differentiation with respect to c, t, or p_3 always means a compensating variation in $p_1 + p_2$. According to these derivatives, differences between various banks' rates of currency spillover will probably have a more significant effect on their loss coefficient than differences in $(r^i + w_2)$. Very likely also, differences in c are more significant than corresponding differences in t. Variations between banks in the magnitude $r^d + w_1$ will produce significant differences in λ only if f_1 or p_3 is sufficiently small. Whenever f_1 and p_3 are large, differences (or changes) in reserve requirements have no appreciable effect on the loss coefficient. In a system with a high degree of decentralization, reflected by a large p_3, and a relatively small rate of utilization of correspondent services, expressed by a large f_1, variations in c are more important than corresponding variations in r^d. Such a system would also exhibit a significant difference in the relevance of r^d between banks in financial centers with relatively low f_1 and banks outside such centers with relatively small f_1. Variations in r^d would have a greater importance for banks in financial centers. Differences in p_3 are not necessarily associated with similar differences in λ. Different levels of p_3 have no effect whatsoever on λ if $f_3 = 1$. Variations in p_3 have the largest impact on λ if $f_3 = 1$. In general, differences in p_3 will induce significant differences in λ provided $f_1 + f_3$ is sufficiently large.

The above statements reveal an indirect significance of the check collection patterns described by the f_i. We consider now their direct effect on λ. Suppose that a larger part of checks redeposited elsewhere are collected via the Federal Reserve mechanism. Thus, we have $df_1 > 0$. This information is not sufficient to specify the effect on λ. Let us, for the moment, consider two alternative specifications which would complete the required information:

i. $$df_1 = -df_3 .$$

In this case, we obtain

$$d\lambda = -(r^d + w_1)p_3 df_1 < 0 .$$

ii. $$df_1 = -df_3 .$$

In this case, we obtain

$$d\lambda = p_3(1 - r^d - w_1)df_1 > 0 .$$

In the first case, the loss coefficient decreases, and in the second case it increases. Both of these opposite movements of λ are associated with an increase in the portion of the checks whose collection involves

a transfer of Federal Reserve funds. The actual outcome of such an increase, denoted by $df_1 > 0$, is thus crucially affected by the compensating changes df_2 and df_3. We can read from the above formulas the effect on λ of $df_1 > 0$, combined with any compensating change $df_2 + df_3 = -df_1$. The qualitative aspects of the results can be arranged in the following table:

TABLE 1a

$d\lambda$	+	0	−
df_2	$> -(1 - r^d - w_1)df_1$ $(\equiv A_1)$	$= A_1$	$< A_1$
df_3	$< -(r^d + w_1)df_1$ $(\equiv A_2)$	$= A_2$	$> A_2$

The same table also holds in case $df_1 < 0$. The rows 2 and 3 indicate the restrictions imposed on the compensating variations df_2 and df_3 in order to assure the sign of $d\lambda$ characterizing a particular column.

A similar table can be derived which summarizes essential features of the case where $df_2 \neq 0$ and $df_3 = -df_1 - df_2$.

TABLE 2a

$d\lambda$	+	0	−
df_1	$> -\frac{df_2}{1 - r^d - w_1}$ $(\equiv A_3)$	$= A_3$	$< A_3$
df_3	$< \frac{r^d + w_1}{1 - r^d - w_1}df_2$ $(\equiv A_4)$	$= A_4$	$> A_4$

This second table exhibits clearly the fact that differences in the proportion of checks collected by debiting the expanding bank's balances at other banks have in themselves no determinate effect on the level of λ. The precise effect depends on the balancing combination of df_1 and df_2.

A table associated with $df_3 \neq 0$ and $df_3 = -df_1 - df_2$ could be easily constructed. The first row of this table would repeat the information already contained in row 3 of Table 1a, and the second row would repeat row 3 in Table 2a. Inspection of the two tables reveals that specific order of magnitude relations determine the meaning of differences in check collection patterns for λ.

APPENDIX II: SOME PROPERTIES OF THE MONETARY MULTIPLIER

The monetary multiplier $(\lambda - \mu)^{-1}$ was derived as a measure of the system's "built-in" magnifying power. This appendix investigates the institutional patterns of the monetary system associated with a relatively large or small multiplier. The results of this investigation

possess a more general interest than the specific higher level hypothesis underlying equation (10) would suggest. The alternative and more general higher level hypothesis mentioned in footnote 11 and constructed in the paper referred to in footnote 3 assigns properties to the multiplier consistent with the patterns implied by the relatively simple hypothesis presented in this paper.

The following table describes the relative change in the multiplier associated with a unit change in each one of its component magnitudes:

TABLE 1b

Unit Change in	Associated with Relative Change of $(\lambda - \mu)^{-1}$
p_3	$-\frac{(f_2 - f_3) + c_2(1 - r^d - w_1) + t_2(r^d + w_2 - r^d - w_1)}{\lambda - \mu}$
r^d and w_1	$-\frac{1 - (c_1 + c_2 p_3) - (t_1 + t_2 p_3)}{\lambda - \mu} < 0$
$c_1 + c_2 p_3$	$\frac{(r^d + w_1 - 1)}{\lambda - \mu} < 0$
$t_1 + t_2 p_3$	$\frac{r^d + w_1 - r^d - w_2}{\lambda - \mu} > 0$
r^d and w_2	$-\frac{t_1 + t_2 p_3}{\lambda - \mu} < 0$
f_2, with f_3 constant	$-\frac{p_3}{\lambda - \mu} < 0$
f_3, with f_2 constant	$\frac{p_3}{\lambda - \mu} > 0$
f_2, with f_1 constant	$-\frac{2p_3}{\lambda - \mu} < 0$
f_3, with f_1 constant	$\frac{2p_3}{\lambda - \mu} > 0$

Variations in r^d have the expected effect on the multiplier; an increase in r^d reduces $(\lambda - \mu)^{-1}$. A rough estimate of the order of magnitude can be obtained in the following manner: Table 2c in Appendix III indicates that $(\lambda - \mu)^{-1}$ is approximately 3.2, and the numerator of the relative change in $(\lambda - \mu)^{-1}$ with respect to r^d will most probably be about .7 or .8. Consequently, the value of the formula approximates 2.5. Multiplication of this magnitude with $100 \times dr^d$ expresses the percentage change in $(\lambda - \mu)^{-1}$ associated with a change in reserve requirements dr^d. Thus, for every .01 point change in r^d, the monetary multiplier will change by about 2.5 percent.

Variations in r^t are most probably of considerably smaller significance to the monetary multiplier. The numerator in the appropriate formula,

i.e., $(t_1 + t_2p_3)$, is at most .15, and thus the reduction of the multiplier associated with a .01 point rise of r^t would be around .5 percent. Thus, the monetary multiplier appears to be insensitive to variations in reserve requirements against time deposits.

An increase in currency spillover rates compresses the system's multiplying power, and an increase in the spillover to time deposits raises it. Comparison of the numerators of the formulas related to c and t suggest that the effect of variations in the currency spillover will dominate the effect of variations in the spillover to time deposits. Thus, differences in the public's marginal propensity to hold currency are more significant for the multiplier than are differences in the public's marginal propensity to hold time deposits.

The re-deposit ratio p_3 may be interpreted as a measure of the banking system's degree of decentralization. The sign of the relative change in $(\lambda - \mu)^{-1}$ induced by a unit change in p_3 is not uniquely determined without further specification. The second term of the numerator has most likely a larger numerical value than the third term; thus the sign of the numerator hinges crucially on the check collection patterns formalized by f_2 and f_3. If a sufficiently large portion of checks are collected by crediting the receiving banks' deposits at issuing banks, then a lower degree of decentralization will compress the multiplier. On the other hand, a sufficiently large portion, f_3, of checks collected by debiting the issuing banks' deposits at receiving banks means that a lower degree of decentralization will raise the multiplier. Also, a sufficiently large portion, f_1, of checks collected by a transfer of Federal Reserve funds will lower $(f_3 - f_2)$ below the positive sum of the second and third terms of the numerator and will thus assure a negative association of p_3 and $(\lambda - \mu)^{-1}$. Whatever the sign of this association may be, an inspection of the formula suggests that it is most likely to be small.

The last four formulas of the table describe the influence of check collection patterns on the monetary multiplier. An increase in the portion, f_2, of checks collected by operating on receiving banks' deposits at issuing banks raises the multiplier, irrespective of whether this increase occurs against a decline in f_1 or f_3. An increase in the portion, f_3, of checks collected by operating on issuing banks' deposits at receiving banks lowers the multiplier, again irrespective of the nature of the compensating change.

On the other hand, the effect of an increase in the portion, f_1, of checks collected by a transfer of Federal Reserve funds depends on the nature of the compensating change. If f_1 increases at the cost

of f_2, then the multiplier increases, and if f_1 increases at the cost of f_2, then the multiplier falls.

APPENDIX III: CORRELATIONS AND REGRESSIONS BEARING ON THE MONEY SUPPLY FUNCTION

TABLE 1c

SIMPLE AND MULTIPLE CORRELATION COEFFICIENTS RELATED TO THE MONEY SUPPLY FUNCTION

Sample Period and Number of Observations	Type of Hypothesis			
	1	2	3	4
2/32-4/37; 23	.98	.92	.98	.998
2/43-4/45; 11	.998	.99	.998	1.00
1/46-4/50; 20	.20	.78	.84	.98
1/51-4/57; 28	.71	.98	.98	.99

The first column in Table 1c states the simple correlations between money stock M and monetary base B. The second column consists of simple correlations between M and $B + L^1$, where L^1 measures the cumulated sum of reserves liberated by changes in average reserve requirements. Column 3 presents simple correlations between M and $B^a + L^1$, where B^a designates the monetary base minus excess reserves. Column 4 consists of multiple correlation coefficients associated with the regression

$$M = b_0 + b_1 c_0 + b_2 t_0 + b_3 (B^a + L^1) .$$

This equation is obtained from equation (14) in the text by deleting the component $m_t \varepsilon_0$ which has been shown to be of negligible relevance and by identifying the magnitude $(w_0 - \gamma_0)$ with the banks' excess reserves.

M = demand deposits adjusted plus currency outside banks. The data used in the first period are quarterly averages of seasonally adjusted monthly values to be published in a forthcoming book on monetary problems by the National Bureau of Economic Research. The data for the other periods are quarterly averages of seasonally unadjusted monthly values obtained from the *Federal Reserve Bulletin*.

B = currency in circulation plus member banks' total reserves. Seasonally unadjusted monthly data were obtained for 1/32–4/37 from *Banking and Monetary Statistics* (Board of Governors of the Federal Reserve System, Washington, 1943) and for the other periods from the *Federal Reserve Bulletin*, table on "Member

Bank Reserves, Reserve Bank Credit and Related Items." The data for the first period were seasonally adjusted by a moving average combined with a free hand adjustment. Quarterly averages were used in the correlations.

B^a = monetary base minus excess reserves. Again, quarterly averages were used.

L^3 = liberated reserves. This magnitude has been computed with the aid of the following formulas:

$$L_t^3 = \sum_{s=1}^{t} l_s^3 ,$$

where

$$l_s^3 = -(r_s^d - r_{s-1}^d)\frac{D_s^n + D_{s-1}^n}{2} - (r_s^t - r_{s-1}^t)\frac{T_s + T_{s-1}}{2} .$$

r_s^d is the average reserve requirement against demand deposits in quarter s; D_s^n are the net demand deposits in quarter s; r_s^t designates average reserve requirements against time deposits in quarter s, and T_s are time deposits in the same quarter.

The table reveals that a simple hypothesis which relates M and B alone cannot be expected to perform adequately. Inclusion of L^3 improves the hypothesis substantially and introduction of $(w_0 - p_0)$ raises the correlation coefficient slightly in two periods. Even so, nearly 30 percent of the total variation observed in M during the third period cannot be explained by the behavior of $(B^a + L^3)$. The multiple correlation coefficient suggests that at least one of the magnitudes c_0 or t_0 contributed markedly in the early post war period to the explanation of the behavior of money stock.

The inadequate predictive power of hypotheses which attempt to explain the money stock in terms of the monetary base only is also revealed by the instability of the resulting multiplier estimates. Columns 1 in Table 2c under M^1 and M^2 show an extreme variability of these estimates. The other columns, associated with extended hypotheses acknowledging the operation of c_0, t_0, and L^3, exhibit a pattern of substantially more stable multiplier estimates.

The column under each measure M^4 of the money stock consists of estimates obtained with the aid of a regression indicated by the appropriate number.

(1c) $$M^4 = a_{04} + a_{14}B ,$$

(2c) $$M^4 = a_{04} + a_{14}B + a_{24}k + a_{34}t + a_{44}r ,$$

(3c) $$M^4 = a_{04} + a_{14}(B + L^3) + a_{24}k + a_{34}t ,$$

(4c) $$M^1 = a_{01} + a_{11}(B^a + L^b) + a_{31}k + a_{41}t\ ,$$

(5c) $$M^1 = a_0 + a_1(B + L^b) + a_2c_0 + a_3t_0\ ,$$

(6c) $$M^1 = a_0 + a_1(B^a + L^b) + a_2c_0 + a_3t_0\ ,$$

$$i = 1, 2\ .$$

M^1 is the sum of demand deposits adjusted plus currency outside banks. M^2 is equal to M^1 plus time deposits of commercial banks. B, B^a, and L^b have already been defined under Table 1c. The definitions of r, the banks' reserve ratio; k, the public's currency ratio; and t, the public's time deposit ratio are

$$r = \frac{C^b + R}{D^p + T}\ ; \quad k = \frac{C^p}{D^p}\ ; \quad t = \frac{T}{D^p}\ ;$$

where C^b = currency held by banks (measured by total currency in circulation minus currency outside banks);

C^p = currency outside banks;

D^p = demand deposits adjusted;

T = time deposits of commercial banks;

R = member banks' total reserves.

Seasonally adjusted monthly data were obtained for 1/32–4/37 from the forthcoming study of the National Bureau of Economic Research for

TABLE 2c

MULTIPLIER ESTIMATES

Sample Period and Number of Observations*	Estimates Associated with									
	M^1						M^2			
	1	2	3	4	5	6	1	2	3	4
1/32-4/37; 24	1·96	2.41 (.04)	1.72 (.88)	3.44 (.14)	1.71 (.84)	3.15 (.12)	2.27	3.43 (.05)	2.42 (1.28)	4.98 (.19)
1/43-4/57; 60	2.90	2.28 (.04)	2.62 (.05)	2.57 (.05)	2.80 (.05)	3.03 (.10)	4.52	2.81 (.07)	3.26 (.09)	3.20 (.09)
1/43-4/45; 12	2.30	2.38 (.03)	2.60 (.09)	2.54 (.07)	2.87 (.19)	3.05 (.26)	3.14	2.99 (.04)	3.28 (.12)	3.20 (.09)
1/46-4/50; 20	.41	2.33 (.09)	2.32 (.34)	2.39 (.28)	2.31 (.34)	2.85 (.16)	.001	3.02 (.11)	3.01 (.43)	3.10 (.35)
1/51-4/57; 28	3.94	2.52 (.09)	2.69 (.14)	2.68 (.15)	2.97 (.15)	4.14 (.23)	6.73	3.32 (.14)	3.53 (.20)	3.53 (.22)

* The first three sample periods for columns 3, 4, 5, and 6 under M^1 and for columns 3 and 4 under M^2 are 2/32-4/37, 2/43-4/57, and 2/43-1/46. The sample periods contain in these cases 23, 59, and 12 observations. The lower number in each cell is the estimate of the standard deviation of the regression coefficient.

$M^1, M^2, C^p, C^p, R, D^p$, and T. For the other periods, seasonally unadjusted monthly data were obtained from the *Federal Reserve Bulletin*. Quarterly averages of monthly data were used in the regressions.

Estimates of equations 5c and 6c which provide also the multiple correlations in Table 1c are based on a preliminary investigation of this type of money supply function. Eventually, demand functions for C^p and T must be appraised and values for c_0 and t_0 derived in this manner. In order to obtain a tentative idea about the resulting money supply function, values for c_0 and t_0 were computed with the aid of an arbitrary assignment defined by $c_0 = C^p - .15M^2$ and $t_0 = T - .15M^2$. Some important results emerging from this trial and error procedure are discussed in the paper mentioned in footnote 3.

An inspection of the table yields the following results:

(a) The estimates in column 2 under M^1 cluster closely together. The 95 percent confidence intervals of all estimates overlap. The difference between the maximal and minimal element is 9.5 percent of the maximal element. The same difference under column 1 (for M^1) measures approximately 90 percent of the maximal estimate.

(b) The estimates of column 2 under M^2 exceed, as expected, the estimates of the corresponding column under M^1. The cluster of estimates is also looser; the 95 percent confidence intervals do not overlap. The difference between maximal and minimal estimate is 12.8 percent of the maximal value. The similar difference in column 1 (for M^2) is approximately 100 percent of the appropriate maximal element.

(c) With the exception of the first sample period columns 3 and 4 under M^1 show a close cluster of estimates in the sense of overlapping confidence intervals. The discrepancy in the period 2/32-4/37 suggests that the estimate 1.72 underestimates the multiplier value and that 3.44 exaggerates its value. This can be attributed to the following circumstance:

The hypothesis constructed is a hyperplane in 5-dimensional space. The regression estimates in column 3 have been derived from an equation which neglects $(w_0 - \gamma_0)$. This regression describes a hyperplane in 4-dimensional space. In this space, the true hypothesis postulated appears as a family of hyperplanes, where each hyperplane is characterized by a fixed and distinct value of $(w_0 - \gamma_0)$. This value contributes to the determination of the M-intercept of the 4-dimensional hyperplane expressing our full hypothesis. The sample period is characterized by increasing

values for both $(w_0 - \gamma_0)$ and $(B + L^3)$. This means that successively higher values of $(B + L^3)$ are located on hyperplanes (in 4 dimensions) with successively lower intercepts. The hyperplane thus generated by the observations consequently slants across the family of hyperplanes expressing the full hypothesis in a manner which lowers its slope with respect to $(B + L^3)$ below the corresponding slope of the family of hyperplanes.

The regression estimates in column 4 (for M^1) were derived from an equation which adds a component to our hypothesis. Excess reserves consist of two viz., (parts, $w_0 - \gamma_0$) and e, the amount associated with deposit liabilities. Our hypothesis contains the magnitude $[B + L^3 - (w_0 - \gamma_0)]$. This magnitude is replaced in equation 4c by $(B^a + L^3)$, and the former magnitudes exceed the latter by e. This circumstance means that the intercept of equation 4c is partly determined by e. The original hypothesis is thus again represented by a family of planes in 4-dimensional space, where each plane is completely specified by a value of e. Both $(B^a + L^3)$ and e increased over the sample periods. This implies that successively larger values of $(B + L^3)$ are associated with hyperplanes characterized by successively higher M-intercepts. The observations thus generate a regression plane which cuts across the family of planes and determine a slope in excess of the corresponding slope of the family of planes. This exaggeration of the slope will be the greater, the greater the increase in e relative to the increase in $(B^a + L^3)$.

(d) The results presented in columns 3 and 4 under M^2 exhibit the same pattern as the corresponding columns under M^1.

(e) Columns 5 and 6 under M^1 again repeat the pattern obtained under columns 3 and 4. The discrepancy observed in the last sample period is particularly intriguing. Over this sample period e and $(w_0 - \gamma_0)$ were most likely moving in opposite directions: e up and $(w_0 - \gamma_0)$ down. Such movements, combined with an increase in $(B + L^3)$ or $(B^a + L^3)$ over the period, resulted in estimates of the slope coefficients which tend to exaggerate its value. This exaggeration is apparently more pronounced in the case of $(B^a + L^3)$ and poses a puzzle to be investigated as the behavior of excess reserves indicates that the changes in $(w_0 - \gamma_0)$ exceeded the simultaneous changes in e.

As for Table 3c, data and sources have already been described. Estimation was made by standard least squares methods. The numbers below the coefficients are estimates of the standard errors of the

TABLE 3c

SOME REGRESSION ESTIMATES OF THE MONEY SUPPLY FUNCTION*

1. $M^1 = 119.55 + \underset{(.04)}{2.28}B - \underset{(4.75)}{146.70}k - \underset{(5.96)}{31.39}t - \underset{(12.98)}{332.97}r$

2. $M^1 = 26.42 + \underset{(.05)}{2.62}(B + L^3) - \underset{(5.72)}{105.04}k - \underset{(5.77)}{5.94}t$

3. $M^1 = 31.32 + \underset{(.05)}{2.57}(B^a + L^3) - \underset{(5.84)}{102.22}k - \underset{(5.93)}{8.94}t$

4. $M^1 = -5.87 + \underset{(.10)}{3.03}(B^a + L^3) - \underset{(.20)}{2.54}c_0 - \underset{(.14)}{.56}t_0$

* The regressions presented correspond to Equations 2, 3, 4, and 6 of Table 2c. In each case the estimate is based on 59 observations in the period 2/43–4/57.

regression coefficients. The difference in the order of magnitude of the coefficients associated with k, t, and r on the one hand, and the multiplier and the coefficients of c_0 and t_0 on the other, has been explained in the paper referred to in footnote 3. The ratios k, t, and r can be interpreted as indexes of the intended magnitudes c_0, t_0 and L^3. These indexes differ from their intended magnitude approximately by the volume of demand deposits adjusted. This scale difference is reflected in the coefficient estimates. It should be noted that all coefficient estimates have signs consistent with our theoretical expectations. Also, for every regression the order of magnitude relation of the k and t, or c_0 and t_0 coefficients conforms to the theory constructed.

INTERNATIONAL ECONOMIC REVIEW
Vol. 2, No. 1, January, 1961

PROFIT, OUTPUT AND LIQUIDITY IN THE THEORY OF FIXED INVESTMENT*

By Murray Brown†

Although many investment theories and econometric forms exist in the literature, there are essentially two great camps in this area; on the one hand, profits as a key explanatory variable and, on the other hand, some form of an output variable compete as determinants of the level of new investment in plant and equipment. These alternative approaches have existed side by side in the same literature. But they exist side by side in the same literature which contains the neo-classical theory of the firm. We raise the inevitable question: Does the neo-classical theory of the firm generate different theories of investment behavior, i.e., can profit- and output-investment theories sprout from the same theoretical foundation? This Medusa-type problem is compounded when we examine the econometric studies of fixed investment. For here various investigators have developed their own "intuitively meaningful" estimating forms, usually based in one or the other great camps of investment theory.

Thus, there are two points of focus which merge in the present paper: (1) we wish to examine the conditions under which profit- and output-investment demand theories are engendered by the neo-classical theory of the firm; and (2) we shall spell out some of the conditions which underlie certain estimating forms that have appeared in the recent econometric literature. From this it may be possible to infer which conditions would generate an empirically superior profits- investment form and which conditions would make reliance upon the output-investment form preferable. It is clear that we are *not* primarily offering a theory of investment, but are trying to spell out the theoretical foundations of existing econometric estimating forms.

In order to accomplish these aims, our analysis assumes the existence of a firm in a government-free market (no taxes) which attempts to minimize a cost function and/or maximize a profits function (revenue minus costs) subject to a production restraint (*inter alia*).[1] The entre-

* Manuscript received July 1, 1959, revised May 26, 1960.

† The author would like to thank Professors Lawrence Klein, Michio Morishima, Hans Neisser, Adolph Lowe, and Mr. Ronald Bodkin for their valuable suggestions and criticisms.

[1] The assumption of a government-free market can be relaxed within the framework of this analysis. See L. Klein [10].

preneur holds static expectations ("tomorrow's price will be like today's") and is free of subjective uncertainty.[2]

1. A BASIC CAPITAL DEMAND FORM

We shall assume at this point that the entrepreneur of our single-product firm behaves in accordance with the principles of cost minimization. He is concerned with the following variables: C = total costs, X = output of the firm, N = services of labor, K = services of capital from the stock of fixed capital, w = the wage rate, q = the price of capital services, including interest and depreciation charges. The unit cost equation is given by

$$\frac{C}{X} = w\frac{N}{X} + q\frac{K}{X} . \tag{1}$$

A constraint on the cost equation is the production function:

$$\bar{X} = f(N, K) \quad \text{where } \bar{X} \text{ is fixed.} \tag{2}$$

Our immediate problem is to determine how much capital is to be used in the production process.[3]

By the method of Lagrange multipliers, the minimization of the unit cost function (1) subject to the production constraint (2) yields two equations:

$$w - \lambda f_N = 0 , \tag{3a}$$

$$q - \lambda f_K = 0 , \tag{3b}$$

[2] For a discussion of stochastic expectations and the theory of the firm, see L. Hurwicz [7]. The present model is a very simple model of the firm. It can only be justified at this stage of our knowledge by the analytical results it yields, i.e., the criterion we use to evaluate its usefulness is its power in generating various extant investment theories and econometric estimating forms which have received attention in the literature. Thus, in terms of the problems handled in this paper, the increase in marginal analytical results yielded by a more complicated model would be less than the marginal analytical work units involved.

[3] There are two slightly different ways of interpreting (1) and (2), both of which formally reduce to the same thing. We can think of the entrepreneur as minimizing costs in terms of the classical theory of the firm, in which case we can derive the well-known ratios of marginal products to factor prices; or we can think of the entrepreneur as minimizing costs over the whole planning horizon, in which case it would be necessary to introduce a discounting procedure. The ratios of marginal products to factor prices in the latter case would represent the present value of the expected marginal product attributable to each factor. Under perfect competition we should replace each price term with prices discounted to the present period. Thus on the formal level there is no difference between the two interpretations. (See [6], pp. 325-326.) Furthermore, it can be shown that there is no need to assume perfect competition in this minimization process. If monopoly elements were to be considered, a new factor would be introduced —the elasticity of supply, (see [10], pp. 103-104)—but no new variables.

where λ is a Lagrange multiplier, and f_N represents the partial of (2) with respect to N, etc. These, together with the production function (2), determine the demand equations of labor and capital—i.e., we obtain optimal values for the decision variables, which, for capital demand is[4]

$$K = K\left(\frac{w}{q}, \bar{X}\right). \tag{4}$$

Since (2) is a constraint on all output, (4) should hold for all X. Therefore,[5]

$$K = K\left(\frac{w}{q}, X\right). \tag{5}$$

Note that if the production function were homogeneous of degree one the capital demand equation would be:

$$K = Xk\left(\frac{w}{q}\right) \tag{5'}$$

where the multiplier of X is a function of w/q alone.

It is possible to derive an optimal capital demand relation from the neo-classical theory of the firm which is similar to (3) under extended maximization conditions. Thus consider the familiar profit identity:

$$\pi = pX - (wN + qK), \tag{6}$$

where p is the price of output. Given a production constraint

$$X = f(N, K), \tag{7}$$

then (6) can be maximized subject to (7) in two steps, a minimization of costs with respect to N and K, provided output is given, and a supplementary maximization of profits with respect to output. (See [16], p. 88.) The cost minimization yields the capital demand relation

$$K = K\left(\frac{w}{q}, X\right), \tag{5}$$

while we obtain

$$X = X\left(\frac{p}{q}, \frac{w}{q}\right) \tag{8}$$

[4] The second order condition for the solution of (2), (3a) and (3b) can be derived in the usual way. The economic interpretation of sufficient conditions among the general conditions are: (a) the marginal productivity of the factors of production must be positive; (b) the marginal productivity of the factors must be decreasing; and (c) the marginal productivity of capital must increase, *cet. par.*, with an addition to the input of labor services.

[5] See [16], pp. 57-69.

by the maximization of profits. Combining (5) and (8) yields

$$K = K\left[\frac{w}{q}, X\left(\frac{p}{q}, \frac{w}{q}\right)\right]. \tag{9}$$

Thus, with full maximization, the capital demand equation relates the services of the capital stock to the ratios of product and factor prices. If the ratio w/q is constant, capital demand depends on output, which may vary as p/q changes; if in addition p/q is kept constant, then, clearly, K remains at a given level. It should be noted that the maximization of profit subject to a production constraint, homogeneous of degree one, does not lead to a finite, unique optimum output.

2. OUTPUT-INVESTMENT THEORIES

If we denote K_{-1} as the capital stock in the previous period and I as net investment, the identity $K = K_{-1} + I$ is obtained which, by substitution into (5) and (9), yields

$$I = K\left(\frac{w}{q}, X\right) - K_{-1}, \tag{10}$$

and

$$I = K\left[\frac{w}{q}, X\left(\frac{p}{q}, \frac{w}{q}\right)\right] - K_{-1}. \tag{11}$$

If the price ratio w/q is held constant in (10), the following investment demand form can be derived:

$$I = K(\bar{a}, X) - K_{-1}, \tag{12}$$

where $\bar{a}$ is a constant. If all prices are held constant in (11), K is a constant, so we shall treat only equation (12). It is to be noticed that (12) represents a genus of investment functions that has appeared in theoretical and empirical studies. R. Eckaus [4] calls (12) the "velocity principle." R. Goodwin [5] considers (12) in terms of an aggregate business cycle model. H. Chenery [2] and P. Clark [3] use a slight variation of (12) in predicting investment demand in the steel and telephone industries, respectively. However, P. Clark rejects the assumption of static expectations which is implicit in (12) and attempts to determine the time profile of entrepreneurial expectations. Finally, F. Modigliani [13] further refines the investment theory which is based on (12) by including the investment planning period explicitly. Before proceeding we should point out that the use of (12) or some variant thereof as an estimating form implicitly assumes the constancy of w/q and cost minimization.

It is very simple to derive the "naive" accelerator, i.e., investment depending on a fixed proportion of the change in output, from (12). We may start with (5) and the identity $I = K - K_{-1}$; then

$$I = K\left(\frac{w}{q}, X\right) - K\left(\frac{w_{-1}}{q_{-1}}, X_{-1}\right). \tag{14}$$

If the production function is homogeneous, we have

$$I = Xk\left(\frac{w}{q}\right) - X_{-1}k\left(\frac{w_{-1}}{q_{-1}}\right). \tag{15}$$

Clearly, the exponents of the naive accelerator hypothesis assume that $w/q = w_{-1}/q_{-1}$.[6]

The two principal assumptions underlying (15) are that the production function is homogeneous and that w/q remains constant in spite of changes in output. But, with a homogeneous production function, prices *can* remain constant while output changes. Professor Morishima has reminded us that the theory of balanced growth shows that outputs can grow or decay in balance with constant prices. Thus, the original acceleration theorists could justify (in terms of the theory of the firm) their basic proposition that a change in output would bring forth a change in decisions concerning the plant size.[7] It is clear then that the fundamental assumption is the homogeneity of the production function. If the production constraint were non-homogeneous, then the inclusion of the output variable in the investment demand equation would be inconsistent with the assumption of constancy of all prices, but consistent with the assumption of the constancy of w/q only.

Yet, in the actual estimation process, an estimating form based on the assumption of homogeneity would yield unstable parameters; for, by linearizing a neo-classical production function, estimates made during periods when firms were operating at increasing returns would obviously differ from estimates made during periods when firms were operating at increasing costs, or at decreasing returns. This may be one source of instability in parameter estimates using a continuous panel of data.

[6] Hurwicz [7], p. 131, and Klein [10], pp. 114-115, also derive an accelerator; Klein derives the naive accelerator by using a production function of the type $X = AN^{\alpha}K^{\beta}$, where A, α and β are parameters. This, of course, coincides with the present derivation except for the assumption of the constancy of prices (and static expectations) which is indicated by our analysis. These authors did not, however, assume $\alpha + \beta = 1$, i.e., homogeneity of degree one. This result could also be derived from a model of the firm which assumes fixed production coefficients.

[7] For a statement of this problem, see H. Neisser [14], pp. 254-255. Also, H. Rose [15] demonstrates that the naive accelerator may be justified provided that a short run output plan—output depending on product prices, factor prices, *inter alia*—is represented.

3. PROFIT-INVESTMENT THEORIES

The second great camp of investment theories involves investment demand as some function of profits. It is possible to derive these estimating forms from the theory of the firm in at least two ways.

1. Consider the factor demand relations:

$$K = K(w/q, X) \tag{5}$$

and

$$N = N(w/q, X)\,, \tag{17}$$

obtainable by the partial maximization procedure—i.e., by minimizing the unit cost function subject to a production constraint. Alternatively, we could begin with $K = Xk(w/q)$ and $N = Xn(w/q)$, which are obtained by a cost minimization procedure but which assume a homogeneous production constraint.

Combining (5) and (17) with the profit identity (6), we have

$$\frac{\pi}{p} = X\left[1 - \frac{w}{p}\frac{N(w/q, X)}{X} - \frac{q}{p}\frac{K(w/q, X)}{X}\right].$$

Let the marginal profit be defined as

$$Q = 1 - \frac{w}{p}h_N - \frac{q}{p}h_K\,,$$

where h_N is the partial derivative of $N(w/q, X)$ with respect to X and h_K is that of $K(w/q, X)$. It is seen that if $Q \neq 0$, X can be written as $X = f(w/q, p/q, \pi/p)$. Hence it follows from (10) that

$$I = K\left(\frac{w}{q}\,,\ \frac{p}{q}\,,\ \frac{\pi}{p}\right) - K_{-1}\,. \tag{10'}$$

The mechanism permitting us to derive the profit investment form is the profit identity; yet this function is undefined at $Q = 0$.

2. It is possible to derive the profit-investment form in a second way—one which is in concordance with statements expressed by certain econometricians who use the profits form in estimating investment demand. Let us assume as before that the entrepreneur maximizes a profit function or minimizes a cost function subject to a production constraint. But now suppose that the entrepreneur is required to consider alternative means of financing a prospective investment project. He is confronted with two alternatives (not mutually exclusive): He can draw upon his savings—i.e., retained earnings plus depreciation—or he can go "outside" his firm—increase his share capital, long term debt, etc. His savings position may be expressed as

$$L = \psi(\pi_{-m}, \cdots, \pi_{-1}, \pi) + D \tag{18}$$

where $\psi(\pi_{-m}, \cdots, \pi_{-1}, \pi)$ represents his retained earnings from $-m$ periods to the present and D is the depreciation component, which for simplicity we shall consider given. If D is somewhat constant through time, then current profits account for the variability in L.

$$L = \psi(\pi_{-m}, \cdots, \pi_{-1}, \pi) + D \equiv \psi^*(L^*, \pi) \ , \tag{19}$$

where $L^* = (\pi_{-m}, \cdots, \pi_{-1})$ and ψ^* includes a constant, D. We can denote the "outside" sources of financing his project by y. The entrepreneur's *investment funds schedule* is

$$\bar{I} = \varphi\{\psi^*(L^*, \pi), y\} \ . \tag{20}$$

What weight will the entrepreneur place on these alternatives? Here, Kalecki's theory of "increasing risk," [8], pp. 91–108, suggests that:

> The expansion of the firm depends on its accumulation of capital out of current profits. This will enable the firm to undertake new investment without encountering the obstacles of the limited capital market or "increasing risk" (i.e., the extent to which the firm risks loss of control either by too great a spread of share capital or by too large an assumption of long time debt). Not only can savings out of current profits be directly invested in the business, but this increase in the firm's capital will make it possible to contract new loans.

The default risk decreases as profits increase, and even if the amount of profits does not change, a decrease in the rate of profits will increase the default risk. Therefore, the willingness and/or ability to finance outside the firm may be written as

$$y = \theta\left(\pi, \frac{\pi}{P_k K_{-1}}\right), \tag{21}$$

where P_k is the price of capital goods. By substitution we have

$$\bar{I} = F\left(L^*, \pi, \frac{\pi}{P_k K_{-1}}\right) .$$

Unless investment funds are obtainable, movements in the variables (as determined by the maximization process) which generate the decisions concerning the size of the firm would be abortive. Hence we are asked to maximize the profit identity, subject to (7) and

[8] The following objection is applicable to (21): the ability to attract external funds in period t is partially a function of the accumulation of past debt as well as present profits, i.e., the larger the past debt in relation to past earnings, the greater is the present "default risk." Thus (21) is deficient in that it does not explicitly identify the role of past debt.

$$P_k I \leqq \bar{I} = F\left(L^*, \pi, \frac{\pi}{P_k K_{-1}}\right). \tag{22}$$

In order to obtain the optimal decision values for X, N, and K, we introduce Lagrange multipliers λ_1 and λ_2 and maximize

$$pX - wN - qK - \lambda_1[X - f(N, K)]$$
$$- \lambda_2\left[P_k(K - K_{-1}) - F\left(L^*, pX - wN - qK, \frac{pX - wN - qK}{P_k K_{-1}}\right)\right].$$

If $\lambda_2 \neq 0$ at the optimal point, the constraint (22) holds with strict equality, so that from (22) we immediately obtain

$$I = \frac{1}{P_k} F\left(L^*, \pi^0, \frac{\pi^0}{P_k K_{-1}}\right), \tag{23}$$

where π^0 stands for the maximal profit. But if $\lambda_2 = 0$ at the optimal point, (22) holds with inequality, i.e., (22) is not an effective constraint, so that the investment function (11) still holds true.

Similar results are obtained, even if we start from cost minimization. If $\lambda_2 \neq 0$ at the point where the unit cost is minimal, we get

$$I = \frac{1}{P_k} F\left(L^*, \pi, \frac{\pi}{P_k K_{-1}}\right). \tag{24}$$

Moreover, if the production function were homogeneous, (24) would hold. If $\lambda_2 = 0$ at that point, a relation similar to (10′) holds true, i.e., the investment function is not affected by the inclusion of the credit restraint.

Let us suppose that F is homogeneous of degree one in L^*, π, and $P_k K_{-1}$. Then (23) and (24) can be written

$$I = G\left(\frac{L^*}{p}, \frac{\pi^0}{p}, K_{-1}, \frac{p}{P_k}\right), \tag{25}$$

$$I = G\left(\frac{L^*}{p}, \frac{\pi}{p}, K_{-1}, \frac{p}{P_k}\right). \tag{25′}$$

If we neglect the effects of L^*/p and p/P_k on I, we obtain from (25) and (25′) the following investment functions:

$$I = g\left(\frac{\pi^0}{p}, K_{-1}\right), \tag{26}$$

$$I = g\left(\frac{\pi}{p}, K_{-1}\right), \tag{26′}$$

respectively. These functions involve two key explanatory variables

used linearly by J. Tinbergen [17], L. Klein [11] and M. Kalecki [8].[9]

4. SUMMARY

We may bring the basic capital demand forms and certain of their assumptions together in Table 1.

Within cell II, the capital demand form (9) is more general (i.e., requires less restrictive assumptions) than (23), which, in turn, is more general than (26). Equations (5′) and (15) in cell III are on the same level of generality; since (24) requires no restriction on the production function but does require $\lambda_2 \neq 0$, it is difficult to say whether it is more or less general than (5′).

In cell IV the most general capital demand equation is (5), since it does not assume constancy of product price. The next least restrictive form is (10′), which does not assume price constancy but does require that marginal profits do not vanish. Approximately on the same level of generality as (5) and (10′) is (24), which requires that that $\lambda_2 \neq 0$, i.e., that the liquidity constraint "swamps" the impact of other variables. Next in the order of generality would come (12), since factor prices are assumed constant. Finally, (26′) is the most restrictive form in cell IV because it assume $\lambda_2 \neq 0$ as well as constancy of p/P_k and of beginning-of-period liquidity.

[9] In one of Kalecki's special cases ([8], p. 100) of his general investment demand equation, the profits variable appears in the form of a change in profits:

$$I = b\left(\Delta \frac{\pi}{p} \Big/ \Delta t\right) + d, \tag{27}$$

where t is a time period, and b and d are parameters. If we assume that $g(\pi/p, K_{-1})$ is of the form $g^*(\pi/p) - K_{-1}$, then we can easily derive

$$I = g^*\left(\frac{\pi}{p}\right) - g^*\left(\frac{\pi_{-1}}{p_{-1}}\right)$$

from (26) or (26′). Furthermore, if g^* is linear, we have

$$I = b\left(\frac{\pi}{p} - \frac{\pi_{-1}}{p_{-1}}\right) \qquad (b = \text{constant}) \tag{28}$$

which is similar to (27) and can be called the "profits accelerator." If g^* is nonlinear, then we cannot write the profit accelerator in its simple form (28); by the theorem of mean value, we would write

$$I = \left(\frac{\pi}{p} - \frac{\pi_{-1}}{p_{-1}}\right) g'\left(\frac{\bar{\pi}}{\bar{p}}\right), \tag{29}$$

where $(\bar{\pi}/\bar{p})$ lies between π/p and π_{-1}/p_{-1}, and g' is the first derivative of the function g^*. Of course, the more general "profits accelerator" is (29).

We may add an empirical note. Although Kalecki found the profits change variable to be significant in his time series investment data, Brown and Roseman [1] found that the simple profits accelerator (28) had no empirical support in their SEC-OBE cross section data

TABLE 1

	Homogeneous Production Function	Non-homogeneous Production Function
	I	II
Profit Maximization	Conditions do not exist for an optimum.	(26) $I = g\left(\frac{\pi^0}{p}, K_{-1}\right)$, (23) $I = \frac{1}{P_k} F\left(L^*, \pi^0, \frac{\pi^0}{P_k K_{-1}}\right)$, (9) $K = K\left[\frac{w}{q}, X\left(\frac{p}{q}, \frac{w}{q}\right)\right]$.
	III	IV
Unit Cost Minimization	(24) $I = \frac{1}{P_k} F\left(L^*, \pi, \frac{\pi}{P_k K_{-1}}\right)$, (5′) $K = Xk\left(\frac{w}{q}\right)$, (15) $I = Xk\left(\frac{w}{q}\right) - X_{-1}k\left(\frac{w_{-1}}{q_{-1}}\right)$.	(5) $K = K\left(\frac{w}{q}, X\right)$, (12) $I = K(\bar{a}, X) - K_{-1}$, (24) $I = \frac{1}{P_k} F\left(L^*, \pi, \frac{\pi}{P_k K_{-1}}\right)$, (10′) $I = K\left(\frac{w}{q}, \frac{p}{q}, \frac{\pi}{p}\right) - K_{-1}$, (26′) $I = g\left(\frac{\pi}{p}, K_{-1}\right)$.

5. PROFITS VERSUS OUTPUT: SOME COMMENTS ON THE CONTROVERSY

If we ignore the investment funds constraint for the moment, it is clear that an output-investment form can be transformed into a profits investment form via the profit identity (revenue minus costs), but since the profit form requires that the marginal profit does not vanish, the output form is more general.

When we include the liquidity constraint in our analysis additional differences arise in the relative effectiveness of the profits and output investment forms. If the liquidity constraint is operative, then a profits form is applicable. Therefore, we can propose three classes of reasons for discrepancies between profits and output forms.

The *first* recognizes that the relative effectiveness of the two forms may depend on the existence of marginal profit.

The *second* involves the investment funds constraint; if this is important in the maximization process, the profit form is more reliable as a predictor of investment activity.

The *third* class of reasons for the discrepancy involves the presence

of rigidities which prohibit the full influence of a profit or output variable to be felt on investment.

An unusually fine example of the effects of "rigidities" on investment decisions is pointed up in a recent econometric study by A. Kisselgoff and F. Modigliani [9] of private investment in the electric power industry. Here, the institutional setting is decisive: While returns are regulated, the utilities are obligated to meet expanded demand; thus a "pressure on capacity variable"—or a form of output variable—is warranted. Since our model is an equilibrium model, we can infer that the institutional rigidities prevented the free working of the profit variable in the determination of investment in this industry.

It may be argued that this inference is neither necessary nor sufficient to account for discrepancies between the results of output and profits investment theories for industry in general, that, alternatively, the profit variable is inferior (superior) to an output variable as a "proxy" for anticipated returns. Of course, this expectational argument would require empirical support that would be difficult to obtain. Our hypothesis offers an explanation of the discrepancies between the results of output and profits investment theories, which is perhaps more amenable to empirical confrontation.

6. AN ECONOMETRIC CONSIDERATION

There are, of course, additional ways of deriving the investment forms discussed above, even if we ignore the "intuitive" or implicit theoretical methods used by some investigators. One important consideration involves statistical criteria in choosing estimating forms. If, for example, we were to focus on macro-relations using firm theory as an analog of the whole economy, then there is no "ultimate" reduced form and consequently no truly independent variables. In this case, statistical criteria are required to choose the variables and form of the investment demand relation. (See [12], pp. 18–21, for a discussion of this point.) The choice of the structural equation would depend on the probability characteristics of each equation and primarily on the stability of the parameters. This criterion may conflict with the theoretical criteria used above.

University of Pennsylvania, U.S.A.

REFERENCES

[1] BROWN, M., AND H. ROSEMAN, "A Cross-Section Analysis of Manufacturing Investment During 1951-1955," *Proceedings of the American Statistical Association* (1957), pp. 344-351.

[2] CHENERY, H., "Overcapacity and the Accelerator," *Econometrica* (January, 1952), pp. 1-28.

[3] CLARK, P., "The Telephone Industry: A Study in Private Investment" in W. W. Leontief *et al.*, *Studies in the Structure of the American Economy*, pp. 243-294. New York: Oxford Press, 1952.

[4] ECKAUS, R. S., "The Accelerator Principle Reconsidered," *Quarterly Journal of Economics* (May, 1954), pp. 209-230.

[5] GOODWIN, R., "Nonlinear Accelerator and Business Cycles," *Econometrica* (January, 1951), pp. 1-17.

[6] HICKS, J. R., *Value and Capital*. Oxford: Clarendon Press, 1939.

[7] HURWICZ, L., "Theory of the Firm and of Investment," *Econometrica* (April, 1946), pp. 109-136.

[8] KALECKI, M., *Theory of Economic Dynamics*. London: Allen and Unwin, 1954.

[9] KISSELGOFF, A., AND F. MODIGLIANI, "Private Investment in the Electric Power Industry and the Acceleration Principle," *The Review of Economics and Statistics* (November, 1957), pp. 363-379.

[10] KLEIN, L., "Notes on the Theory of Investment," *Kyklos*, Vol. II (1948), pp. 97-117.

[11] ———, "Studies in Investment Behavior," *Conference on Business Cycles*, pp. 233-277. New York: National Bureau of Economic Research, 1951.

[12] ———, *Textbook of Econometrics*. Evanston: Row, Peterson, 1953.

[13] MODIGLIANI, F., "Comments on B. C. Hickman," "Capacity, Capacity Utilization and the Acceleration Principle," in *Problems in Capital Formation*, National Bureau of Economic Research, Vol. 19 (1957), pp. 450-462.

[14] NEISSER, H., "Critical Notes on the Accelerator," *Quarterly Journal of Economics* (May, 1954), pp. 253-274.

[15] ROSE, H., "Demand, Supply and the Price Level in Macro-Economics," *The Review of Economic Studies* (1953), pp. 1-23.

[16] SAMUELSON, P., *Foundations of Economic Analysis*. Cambridge: Harvard, 1955.

[17] TINBERGEN, J., *A Method and its Application to Investment Activity*. Geneva, League of Nations, 1939.

INTERNATIONAL ECONOMIC REVIEW
Vol. 2, No. 1, January, 1961

ON THE FORMATION OF PRICES*

By Takashi Negishi[1]

1. INTRODUCTION

In the theories of the formation of prices under competition it is usually assumed that no exchange transaction will be undertaken until a system of prices is reached which secures equilibrium in all parts of the market. The use of tickets in tâtonnement (Walras [8], p. 37) or the possibility of recontract (Edgeworth [2], pp. 311-314) is introduced for this purpose. Because of this assumption, the equilibrium becomes determinate in the sense that the process of adjustment does not alter the conditions of equilibrium (Kaldor [4], pp. 123-127).

In the absence of this assumption, the equilibrium is generally indeterminate and depends on the route followed by prices before the equilibrium is reached. The indeterminateness of the equilibrium arises through redistributions of incomes among individuals due to changes of prices in the midst of trading (Hicks [3], pp. 127-129). In this note we shall discuss the stability of the indeterminate equilibrium. We shall first show the stability under the assumption of gross substitutability. Second, we shall treat a case in which a social utility function exists. Because we cannot specify the equilibrium prices beforehand, the discussion of the stability is necessarily a global one.

2. CONSTRUCTION OF THE MODEL

Let us denote

P_j = price of the j-th commodity, $j = 1, \cdots, m$;
I_i = income of the i-th individual, $i = 1, \cdots, n$;
$\bar{X}_{ij}$ = holding of the j-th commodity by the i-th individual;
$\bar{X}_j$ = total amount of the j-th commodity, constant;
X_{ij} = demand for the j-th commodity by the i-th individual;
X_j = total demand for the j-th commodity;
U_i = utility of the i-th individual;
P = price vector composed of P_j, $j = 1, \cdots, m$;
$\bar{X}$ = a matrix composed of $\bar{X}_{ij}$, $i = 1, \cdots, n, j = 1, \cdots, m$.

Let us define,

* Manuscript received May 3, 1960, revised August 15, 1960.

[1] The author is greatly indebted to Professors K. J. Arrow, F. H. Hahn and H. Uzawa for their valuable suggestions. This work was supported by the Office of Naval Research under Task NR-047-004 at Stanford University.

DEFINITION 1. *The demand for the j-th commodity by the i-th individual,* $X_{ij}(P_1, \cdots, P_n, I_i)$ *is a solution of maximizing* $U_i(X_{i1}, \cdots, X_{im})$ *subject to* $\sum_j P_j X_{ij} = \sum_j P_j \bar{X}_{ij} \equiv I_i$.

DEFINITION 2. *The condition of equilibrium is*

$$X_j(P_1, \cdots, P_m, I_1, \cdots, I_n) \equiv \sum_i X_{ij} = \sum_i \bar{X}_{ij} \equiv \bar{X}_j, \quad \text{for } j = 1, \cdots, m.$$

In Definition 2 we presuppose that the equilibrium prices are all positive. We assume,

ASSUMPTION 1. *By exchange transaction, holding of a commodity becomes a certain continuous differentiable function of time* $\bar{X}_{ij}[t; P(0), \bar{X}(0)]$ *such that* $\sum_i \bar{X}_{ij}(t) = \bar{X}_j$, $\sum_j P_j(d\bar{X}_{ij}/dt) = 0$; *it is continuous with respect to initial conditions,* $P_j(0)$, $\bar{X}_{ij}(0)$, *for all* i, j.[2]

Assumption 1, which implies no credit transactions, assures a positive income for individuals provided that prices and initial holdings are positive. The equilibrium defined in Definition 2 becomes indeterminate and depends on $\bar{X}_{ij}(t)$, except for the case of $\bar{X}_{ij}$ constant through time, which implies the usual assumption that no transaction is undertaken until equilibrium is reached.

ASSUMPTION 2. *The dynamic path of prices is determined by* $\dot{P}_j = X_j - \bar{X}_j$, $j = 1, \cdots, m$, *subject to Walras' identity* $\sum_j P_j(X_j - \bar{X}_j) \equiv 0$, *which is deduced from* $\sum_j P_j X_{ij} = \sum_j P_j \bar{X}_{ij}$ *in Definition* 1. *As the solution of this dynamic system, prices are continuous functions of time* $P_j[t\,; P(0), \bar{X}(0)]$, *such that* $P_j(t) \geqq \delta_j > 0$ *for all* j *and are continuous with respect to initial conditions,* $P_j(0)$, $\bar{X}_{ij}(0)$, *for all* i, j.[3]

3. STABILITY AND GROSS SUBSTITUTABILITY

In this section we assume

ASSUMPTION 3. *All the commodities are strict gross-substitutes, such that*

$$\frac{\partial X_j}{\partial P_k} + \sum_i \frac{\partial X_j}{\partial I_i} \frac{\partial I_i}{\partial P_k} > 0\,, \quad j \neq k\,.$$

[2] It may be considered that $\bar{X}_{ij}(t)$ is generated from

$$\dot{\bar{X}}_{ij} = F_{ij}(P_1, \cdots, P_m, \bar{X}_{11}, \cdots, \bar{X}_{1m}, \bar{X}_{21}, \cdots, \bar{X}_{nm})\,.$$

[3] This can be assured by assuming that there are positive numbers $\gamma_s, \delta_s, \varepsilon_s, (\gamma_s > \delta_s, s = 1, \cdots, m)$ such that $P_j \leqq \delta_j$, $P_k \geqq \delta_k$, $(k \neq j)$, $\sum_s P_s^2 \geqq \sum_s \gamma_s^2$, implies $X_j - \bar{X}_j \geqq \varepsilon_j > 0$ for any distribution of $\bar{X}_{ij}$, and $P_j(0) \geqq \gamma_j$ for all j.

We are now in a position to state

THEOREM 1. *Under Assumptions* 1-3, *the process* $\dot{P}_j = X_j - \bar{X}_j$, $j = 1, \cdots, m$, *is stable in the sense that the excess demands of all commodities* $X_j - \bar{X}_j, j = 1, \cdots, m$, *converge to zero.*

PROOF: Let us construct a characteristic function as follows: $V(t) = \sum_{j \in J^+(t)} P_j(t)[X_j(t) - \bar{X}_j]$, where $J^+(t) = \{j \mid X_j(t) - \bar{X}_j > 0$, or $X_j(t) - \bar{X}_j = 0, \dot{X}_j(t) > 0\}$. By the Walrasian identity, we have $V(t) = \sum_{j \in J^-(t)} [-P_j(t)][X_j(t) - \bar{X}_j]$, where $J^-(t) = \{j \mid X_j(t) - \bar{X}_j < 0$, or $X_j(t) - \bar{X}_j = 0, \dot{X}_j(t) < 0\}$.

This function was first used by Allais [1], Vol. 2, pp. 486-489, and then studied extensively by McKenzie [5].

For sufficiently small $h > 0$, we have

$$V(t + h) = \sum_{j \in J^+(t)} P_j(t + h)[X_j(t + h) - \bar{X}_j] + \sum_{j \in J} P_j(t + h)[X_j(t + h) - \bar{X}_j],$$

where J is a suitable subset of $J^0(t) = \{j \mid X_j(t) - \bar{X}_j = 0, \dot{X}_j(t) = 0\}$.

We wish to show that

$$W(t) = \overline{\lim_{h \to 0}} \frac{1}{h}\left\{V(t + h) - V(t)\right\} = \lim_{v \to \infty} \frac{1}{h_v}\left\{V(t + h_v) - V(t)\right\} \leqq 0,$$

where $\{h_v\}$ is a sequence such that $h_v \to 0$ as $v \to \infty$, and equality holds only at equilibrium.

Because the number of commodities, m, is finite, we can take a suitable subsequence of h_v and find

$$\lim_{v \to \infty} \frac{1}{h_v}\left\{V(t + h_v) - V(t)\right\} = \sum_{j \in J^+(t)} \frac{d\{P_j(X_j - \bar{X}_j)\}}{dt} + \sum_{j \in J} \frac{d\{P_j(X_j - \bar{X}_j)\}}{dt}$$

for a certain $J \subset J^0(t)$.

From the definition of $J^0(t)$, $\sum_{j \in J} d[P_j X_j]/dt = 0$, we have $W(t) = \sum_{j \in J^+(t)} d[P_j X_j]/dt$. From the definition of X_j we have

$$\sum_{j \in J^+(t)} \frac{d[P_j X_j]}{dt} = \sum_k \sum_{j \in J^+(t)} \frac{\partial[P_j X_j]}{\partial P_k} \dot{P}_k + \sum_i \sum_k \sum_{j \in J^+(t)} \frac{\partial[P_j X_j]}{\partial I_i} \frac{\partial I_i}{\partial P_k} \dot{P}_k + \sum_i \sum_k \sum_{j \in J^+(t)} \frac{\partial[P_j X_j]}{\partial I_i} \frac{\partial I_i}{\partial \bar{X}_{ik}} \dot{\bar{X}}_{ik}.$$

The last term of the right hand side of the above equation can be written as

$$\sum_i \sum_{j \in J^+(t)} \frac{\partial[P_j X_j]}{\partial I_i} \sum_k P_k \dot{\bar{X}}_{ik}.$$

which vanishes because of Assumption 1, while by the use of the Walrasian identity the other terms can be expressed as

$$-\sum_{k\in J^+(t)}\left\{\sum_{j\in J^+(t)}\frac{\partial[P_jX_j]}{\partial P_k}+\sum_i\sum_{j\in J^+(t)}\frac{\partial[P_jX_j]}{\partial I_i}\frac{\partial I_i}{\partial P_k}\right\}\dot{P}_k$$

$$+\sum_{k\in J^-(t)}\left\{\sum_{j\in J^+(t)}\frac{\partial[P_jX_j]}{\partial P_k}+\sum_i\sum_{j\in J^+(t)}\frac{\partial[P_jX_j]}{\partial I_i}\frac{\partial I_i}{\partial P_k}\right\}\dot{P}_k\,,$$

which is negative, from Assumption 3, for $j \neq k$ in each term. At equilibrium, J^+ and J^- become null sets and we have $V = 0$, $W = 0$. The domain of $P(t)$ is bounded because we have $d(\sum_j P_j^2)/dt = 2\sum_j P_j\dot{P}_j = 0$. The domain of $\bar{X}(t)$ is also bounded. Since $V(t)$ is bounded below by 0, the fact that $V(t)$ decreases implies the existence of $\lim_{t\to\infty} V(t) = V^*$. Let P_j^*, $\bar{X}_{ij}^*$ be a limit of $P_j(t)$, $\bar{X}_{ij}(t)$ as $t \to \infty$, i.e., $P_j^* = \lim_{k\to\infty} P_j[t_k; P(0), \bar{X}(0)]$, $\bar{X}_{ij}^* = \lim_{k\to\infty} \bar{X}_{ij}[t_k; P(0), \bar{X}(0)]$, where $t_k \to \infty$ as $k \to \infty$. Then we have $V^*(t) = V^*[t; P^*, \bar{X}^*] = \lim_{k\to\infty} V[t; P(t_k), \bar{X}(t_k)] = \lim_{k\to\infty} V[t + t_k; P(0), \bar{X}(0)] = V^*$, from the continuity of P_j, $\bar{X}_{ij}$, and V. Hence, $W(t) = 0$ at P^*, $\bar{X}^*$. We can, therefore, conclude that $X_j - \bar{X}_j \to 0$ as $t \to \infty$. See Uzawa [6].

4. STABILITY AND SOCIAL UTILITY

In this section we make the following assumptions:

ASSUMPTION 4. *The utility function $U_i(X_{i1}, \cdots, X_{im})$ is strictly quasi-concave and homogeneous with respect to all variables.*

ASSUMPTION 5. *Utility functions of all individuals are the same, i.e., $U_i = U$, $i = 1, \cdots, n$.*

Then we have:

THEOREM 2. *Under Assumptions 4 and 5 total demand for the j-th commodity X_j is the solution of maximizing $U(X_1, \cdots, X_m)$, subject to $\sum_j P_jX_j = \sum_j P_j\bar{X}_j$, and the equilibrium prices are determinate and unique.*

PROOF: By definition $X_j = \sum_i X_{ij}$. From Definition 1 and Assumption 4, $X_{ij} = A_{ij}(P_1, \cdots, P_m)\sum_j P_j\bar{X}_{ij}$, and from Assumption 5, $X_{ij} = A_j(P_1, \cdots, P_m)\sum_j P_j\bar{X}_{ij}$. Since $\sum_i \bar{X}_{ij} = \bar{X}_j$,

$$X_j = \sum_i X_{ij} = A_j(P_1, \cdots, P_m)\sum_j P_j\bar{X}_j\,,$$

which is the solution of maximizing $U(X_1, \cdots, X_m)$ subject to $\sum_j P_jX_j = \sum_j P_j\bar{X}_j$ from Assumptions 4 and 5.

Total demand $X_j(P)$ is independent of the distribution of incomes among individuals. The equilibrium price vector $\bar{P}$ such that

$X_j(\bar{P}_1, \cdots, \bar{P}_m) = \bar{X}_j$ in Definition 2 is determinate and independent of the dynamic path of $P(t)$, $\bar{X}(t)$, although the X_{ij}'s are still indeterminate and depend on the route followed by $P(t)$, $\bar{X}(t)$. As a result of strict quasi-concavity, equilibrium prices are unique.

We state now:

THEOREM 3. *Under Assumptions* 1, 2, 4 *and* 5, *the expression* $\dot{P}_j = X_j - \bar{X}_j$, $j = 1, \cdots, m$, *is stable in the sense that prices converge to the equilibrium prices.*[4]

PROOF: By differentiating Walras' identity, $\sum_j \dot{P}_j(X_j - \bar{X}_j) + \sum_j P_j\dot{X}_j = 0$, and from $\dot{P}_j = X_j - \bar{X}_j$, we have $\sum_j P_j\dot{X}_j \leqq 0$, with equality holding only at equilibrium. If we set $\hat{U}(t) = U\{X_1(t), \cdots, X_m(t)\}$ and differentiate $\hat{U}$ with respect to t, we have from Theorem 2 $d\hat{U}/dt = \sum_j (\partial U/\partial X_j)\dot{X}_j = \lambda\sum_j P_j\dot{X}_j \leqq 0$, where λ is a positive function of t. See Allais [1], Vol. 1, p. 26. Because $\hat{U}(t)$ is decreasing and bounded below, and $P(t)$ is bounded, we can show prices converge to the unique equilibrium. See Uzawa [6], [7].

Stanford University, U. S. A., and
Tokyo University, Japan.

REEERENCES

[1] ALLAIS, M., *Traité d'économie pure*. Paris: Imprimerie Nationale, 1952.
[2] EDGEWORTH, F. Y., *Papers Relating to Political Economy*, Vol. 2. London: Macmillan and Co., Ltd., 1925.
[3] HICKS, J. R., *Value and Capital*. Oxford: Oxford University Press, 1946.
[4] KALDOR, N., "A Classificatory Note on the Determinateness of Equilibrium," *The Review of Economic Studies*, Vol. 1 (1933-34), pp. 122-136.
[5] MCKENZIE, L. W., "A Contribution to the Theory of the Stability of Competitive Equilibrium," (abstract) *Econometrica*, Vol. 27 (1959), p. 303.
[6] UZAWA, H., "On the Stability of Dynamic Processes," *Econometrica*, forthcoming.
[7] ———, "The Stability of Competitive Economy," (abstract) *Econometrica*, Vol. 27 (1959), p. 304.
[8] WALRAS, L., *Elements of Pure Economics*, trans. W. Jaffé. Homewood, Illinois: Richard D. Irwin, Inc., 1954.

[4] A similar result was obtained by L. Hurwicz and H. Uzawa, see Uzawa [7].

INTERNATIONAL ECONOMIC REVIEW
Vol. 2, No. 1, January, 1961

PRODUCTION FUNCTIONS WITH RESTRICTED INPUT SHARES*

BY P. K. NEWMAN AND R. C. READ

IT IS COMMON in many economic models to impose restrictions on relative input shares. One such restriction is to suppose that the relative input shares are constant, and it is known that the production function must then be of a generalized Cobb-Douglas type. A proof of this will be given here[1]. In this paper we consider some general and rather less stringent restrictions on the input shares and show that certain other functions, not of a generalized Cobb-Douglas type, are then possible as production functions.

The inputs will be denoted by $x, y, z, \cdots$ and the relative input shares by $a, b, c, \cdots$ respectively, these being, in general, functions of the inputs. We shall consider separately the two following sets of restrictions:

1. The relative share of an input remains invariant with changes in that input itself, the other inputs being held constant.

2. The relative share of one input remains invariant with changes in another input, the remaining inputs being held constant.

We shall deal with each case in turn.

1. INVARIANCE WITH RESPECT TO THE SAME INPUT

We first consider a production function $F(x, y)$ with only two inputs. Then, if factors are rewarded according to their marginal products, the relative shares of x and y, denoted by $a(x, y)$ and $b(x, y)$, are defined by

$$a(x, y) = \frac{xF_x}{F}, \qquad b(x, y) = \frac{yF_y}{F}, \tag{1}$$

where F_x, F_y are the partial derivatives of F with respect to x and y respectively.

It turns out that the computations we shall need to make are easier

* Manuscript received June 1, 1959, revised June 22, 1960.

[1] See also the discussion between W. P. Hogan and R. M. Solow, *Review of Economics and Statistics*, Vol. XL (1958), pp. 407-413, for a proof under the restriction of constant returns to scale.

if we deal with the logarithms of F, x, and y instead of these quantities themselves. Accordingly, we put

$$\begin{aligned} \xi &= \log x\,, \\ \eta &= \log y\,, \\ \varphi &= \log F\,, \end{aligned} \tag{2}$$

and the equations (1) take the simple forms

$$\begin{aligned} a(x, y) &= \frac{\partial \varphi}{\partial \xi}\,, \\ b(x, y) &= \frac{\partial \varphi}{\partial \eta}\,. \end{aligned} \tag{3}$$

Our restrictions in this section imply that $a(x, y)$ depends only on y, and $b(x, y)$ depends only on x. Hence, by (3), $\partial\varphi/\partial\xi$ is a function of η alone, and $\partial\varphi/\partial\eta$ is a function of ξ alone.

Let us write

$$\frac{\partial \varphi}{\partial \xi} = \alpha(\eta)\,, \tag{4}$$

$$\frac{\partial \varphi}{\partial \eta} = \beta(\xi)\,. \tag{5}$$

Integrating (4), we obtain

$$\varphi = \alpha(\eta)\xi + A(\eta)\,, \tag{6}$$

where $A(\eta)$ is a function of η, as yet arbitrary.

From (6) and (5), we obtain

$$\alpha'(\eta)\xi + A'(\eta) = \beta(\xi)\,,$$

from which it follows that $\alpha'(\eta)$ and $A'(\eta)$ are constants, so that $\alpha(\eta)$ and $A(\eta)$ are linear expressions in η. If we put

$$\alpha(\eta) = a + \lambda\eta\,, \tag{7}$$

$$A(\eta) = k + b\eta\,, \tag{8}$$

where k, a, b, λ are constants, we see that

$$\varphi(\xi, \eta) = k + a\xi + b\eta + \lambda\xi\eta\,. \tag{9}$$

In the original variables, we have

$$F(x, y) = Kx^a y^b \exp\{\lambda(\log x)(\log y)\}\,, \tag{10}$$

where $K = e^k$ is a constant.

Thus, as might have been expected, our less stringent restrictions have led us to a more general production function, of which the

Cobb-Douglas function is a special case ($\lambda = 0$).

It is of interest to examine the values of x and y for which the function (10) enjoys some of the properties usually required of production functions, viz. that F_x and F_y are positive, that F_{xx} and F_{yy} are negative, and that the curves, $F(x, y) = \text{constant}$, are concave.

Now,

$$\frac{xF_x}{F} = \alpha(\eta) = a + \lambda\eta$$

by (1), (3), (4) and (7). Hence

$$F_x = \frac{a + \lambda \log y}{x} F , \tag{11}$$

and

$$F_{xx} = \frac{\alpha + \lambda \log y}{x} F_x - \frac{a + \lambda \log y}{x^2} F , \tag{12}$$

$$F_{xx} = \frac{1}{x^2}(a + \lambda \log y)(a + \lambda \log y - 1)F .$$

Since F, x, and y are positive, we see from (11) that $F_x > 0$ if, and only if, $a + \lambda \log y > 0$; and from (12) that $F_{xx} < 0$ if, and only if, $a + \lambda \log y < 1$.

Hence, we must have

$$0 < a + \lambda \log y < 1. \tag{13}$$

Similarly, we must have

$$0 < b + \lambda \log x < 1 \tag{14}$$

for $F_y > 0$ and $F_{yy} < 0$.

The condition for the concavity of the curves $F(x, y) = \text{constant}$ is[2]

$$2F_{xy}F_xF_y - F_x^2F_{yy} - F_y^2F_{xx} > 0 .$$

Writing, as before, $\alpha = a + \lambda \log y$, $\beta = b + \lambda \log x$, and observing that

$$F_{xy} = \frac{1}{xy}(\alpha\beta + \lambda)F ,$$

we see that this condition becomes

$$\frac{2}{x^2y^2}(\alpha\beta + \lambda)\alpha\beta F^3 - \frac{\alpha^2}{x^2}\frac{\beta(\beta - 1)}{y^2}F^3 - \frac{\beta^2}{y^2}\frac{\alpha(\alpha - 1)}{x^2}F^3 > 0 ,$$

[2] See e.g. P. A. Samuelson, *Foundations of Economic Analysis* (Cambridge: Harvard University Press, 1947), p. 62.

which reduces to

$$2\lambda + \alpha + \beta > 0 . \tag{15}$$

If λ is positive, (15) is satisfied, since α and β are positive.

Hence, if λ is positive, the stated properties all hold inside the rectangle

$$\begin{aligned} \exp\left(\frac{-b}{\lambda}\right) < x < \exp\left(\frac{1-b}{\lambda}\right), \\ \exp\left(\frac{-a}{\lambda}\right) < y < \exp\left(\frac{1-a}{\lambda}\right). \end{aligned} \tag{16}$$

If λ is negative, then (13) and (14) give

$$\begin{aligned} \exp\left(\frac{1-b}{\lambda}\right) < x < \exp\left(\frac{-b}{\lambda}\right), \\ \exp\left(\frac{1-a}{\lambda}\right) < y < \exp\left(\frac{-a}{\lambda}\right). \end{aligned} \tag{17}$$

From (15) we must have

$$(2 + \log x + \log y)\lambda > -a - b ,$$

i.e.

$$2 + \log x + \log y < -\frac{a+b}{\lambda} , \qquad \text{since } \lambda < 0 ,$$

or

$$xy < \exp\left(-\frac{a+b}{\lambda} - 2\right). \tag{18}$$

Since the right-hand side of (18) is less than the greatest value, $\exp(-(a+b)/\lambda)$, of xy in the rectangle, it follows that (18) is a further restriction on the values of x and y for which the curves $F(x, y) =$ constant are concave.

It is easily verified that if $\lambda < -1$, the right-hand side of (18) is less than the least value of xy over the rectangle, and the curves, $F(x, y) =$ constant, are not concave for any values of x and y.

The function F gives constant returns to scale only in the original Cobb-Douglas case, where $\lambda = 0$ and $a + b = 1$. Otherwise F may give increasing or decreasing returns to scale depending on the values of the three parameters a, b, and λ.

Generalization of the above result to more than two inputs is straightforward. For three inputs it can be shown that we must have

$$\varphi(\xi, \eta, \zeta) = k + a\xi + b\eta + c\zeta + \lambda\eta\zeta + \mu\zeta\xi + \nu\xi\eta + \rho\xi\eta\zeta , \tag{19}$$

where $k, a, b, c, \lambda, \mu, \nu, \rho$ are constants, and $\zeta = \log z$. The characteristic property of φ is thus that it is linear in each of the variables $\xi, \eta, \zeta, \cdots$.

In the original variables, (19) becomes

$$F(x, y, z) = Kx^a y^b z^c \exp\{\lambda(\log y)(\log z) + \eta(\log z)(\log x) + \nu(\log x)(\log y) + \rho(\log x)(\log y)(\log z)\} . \tag{20}$$

2. INVARIANCE WITH RESPECT TO ANOTHER INPUT

We now consider the second set of restrictions, viz. that the relative share of one input remains invariant with changes in another input, the remaining inputs being held constant.

We shall take the case of a three-input production function $F(x, y, z)$ and examine first the effect of the restriction that x's share, $a(x, y, z)$, defined by xF_x/F, remains invariant with changes in y alone, i.e. that $a(x, y, z)$ is, in fact, a function of x and z only.

The equation analogous to (4) then becomes

$$\frac{\partial \varphi}{\partial \xi} = \alpha(\xi, \zeta) ,$$

whence

$$\varphi = \int \alpha(\xi, \zeta) d\xi + \kappa(\eta, \zeta) ,$$

$$\varphi = \mu(\xi, \zeta) + \kappa(\eta, \zeta) ; \tag{21}$$

i.e. φ is the sum of two functions, one independent of η, the other independent of ξ. In the original variables we have

$$F(x, y, z) = M(x, z)K(y, z) .$$

In words, the production function must factorize into a pair of functions neither of which involves both x and y—a conclusion that is more or less intuitively obvious. It follows also that no variation in x can alter y's share b.

In like manner, the condition that $a(x, y, z)$ remains invariant with changes in z is that

$$\varphi = \rho(\xi, \eta) + \nu(\eta, \zeta) . \tag{22}$$

Hence, if both restrictions apply, we have from (21) and (22) that

$$\mu(\xi, \zeta) = \rho(\xi, \eta) + \nu(\eta, \zeta) - \kappa(\eta, \zeta) , \tag{23}$$

for all ξ, η, ζ.

Putting $\eta = 0$, we have

$$\mu(\xi, \zeta) = \rho(\xi, 0) + [\nu(0, \zeta) - \kappa(0, \zeta)] ,$$

for all ξ, ζ; i.e. $\mu(\xi, \zeta)$ is the sum of a function of ξ alone and a function of ζ alone.

Hence, from (21), φ is the sum of a function $\rho(\xi)$ of ξ alone, and a function of η and ζ.

Thus the result of applying the restriction that x's share is invariant with variations in each of the other variables separately is that the terms in φ which involve ξ "split off" as a function of ξ alone. If we therefore apply also the same restrictions to the shares of y and z, the function must split up into the sum of three functions, one in each of the variables, viz.

$$\varphi = \rho(\xi) + \sigma(\eta) + \gamma(\zeta) . \tag{24}$$

In the original variables (24) becomes

$$F(x, y, z) = R(x)S(y)T(z) , \tag{25}$$

i.e. the production function is completely "factorable" into a product of functions of each variable alone.[3]

3. SUMMARY

Under our first set of restrictions the production function turns out to be the function $Kx^a y^b \exp\{\lambda(\log x)(\log y)\}$, which does not seem to have appeared in the literature before. The generalized Cobb-Douglas function is a special case of this when $\lambda = 0$; if, further, $a + b = 1$ we have the original Cobb-Douglas function.

Under our second set of restrictions the production function may be any product of functions in each of the variables separately. The function given above is not of this form unless $\lambda = 0$; hence we may deduce that if $a, b, c \cdots$ are absolute constants, the production function must be of generalized Cobb-Douglas type. We note, however, that this is as easily proved, *ab initio*, by direct integration of equation (1) with a as a constant, and so on.

In conclusion we mention an interesting interpretation of an industry

[3] The internal structure of such factorable functions has been examined in some detail by W. Leontief, *Econometrica*, Vol. 15 (1947), pp. 361-373. Such functions are obviously relevant to the construction of aggregate production functions from the underlying micro-functions. See R. W. Shephard, *Cost and Production Functions* (Princeton: Princeton University Press), 1953, chapter 9.

Cobb-Douglas function given by Houthakker[4]. If the constituent firms of the industry have linear production functions, and the technical coefficients form a joint probability distribution of the Pareto type, then the aggregate industry production function will be of the generalized Cobb-Douglas type. This seems a promising micro-economic line of attack on the problem of constant input shares[5].

University College of the West Indies, Jamaica

[4] H. S. Houthakker, *Review of Economic Studies*, Vol. XXIII (1955-56), pp. 27-31.

[5] It is also probably a more promising line of attack on the problem of constructing and interpreting aggregate functions than the deterministic approach mentioned in the third footnote of this paper.

ECONOMETRICA

Journal of the Econometric Society

Vol. 28, No. 3 **July, 1960**

CONTENTS

JAMES DUESENBERRY, OTTO ECKSTEIN, and GARY FROMM: Stability and Instability in the American Economy.

A. H. LAND and A. G. DOIG: An Automatic Method of Solving Discrete Programming Problems.

RALPH E. GOMORY and WILLIAM J. BAUMOL: Integer Programming and Pricing.

I. RICHARD SAVAGE and KARL W. DEUTCH: A Statistical Model for the Gross Analysis of Transaction Flows.

A. L. NAGAR: A Monte Carlo Study of Alternative Simultaneous Equation Estimators.

GREGORY C. CHOW: Tests of Equality between Sets of Coefficients in Two Linear Regressions.

LIONEL W. MCKENZIE: A Contribution to the Theory of Competitive Equilibrium.

W. V. CANDLER: A Short-Cut Method for the Complete Solution of Game Theory and Feed-Mix Problems.

CHARLES W. HOWE: An Alternative Proof of the Existence of General Equilibrium in a von Neumann Model.

KENNETH J. ARROW and LEONID HURWICZ: Some Remarks on the Equilibria of Economic Systems.

Published Quarterly Subscription rate available on request

The Econometric Society is an international society for the advancement of economic theory in its relation to statistics and mathematics.

Subscriptions to ECONOMETRICA and inquiries about the work of the Society and the procedure in applying for membership should be addressed to Richard Ruggles, Secretary, The Econometric Society, Box 1264, Yale Station, New Haven, Connecticut, U.S.A.

THE JOURNAL OF POLITICAL ECONOMY

Edited by HARRY G. JOHNSON, in Co-operation with the Other Members of the Economics Department of the University of Chicago

The February 1961 issue contains:

Statistical Tests of Rival Monetary Rules .....*Martin Bronfenbrenner*

Market Practices and Collective Bargaining in Auto Parts ..... *Kenneth Alexander*

Economic Factors in Specie Resumption, the United States, 1865-1879 .....*James K. Kindahl*

Counter Speculation and the Forward Exchange Market.....*John Auten*

Gurley and Shaw on Money in a Theory of Finance .....*Alvin L. Marty*

Volume Discounts as Barriers to Entry and Access ....*Robert C. Brooks, Jr.*

Inflation as a Tax in World War II .....*Ralph Turvey*

Book Reviews and Books Received

Published bimonthly in February, April, June, August, October, and December. Subscription rates for U.S.A. and possessions: 1 year, $6.00; 2 years, $11.50; 3 years, $16.50. Canada and Pan American Postal Union: Fifty cents more per year. All other countries: $1 more per year. Single copies; $1.75. Special Rate: $3.00 per year to full-time graduate students in economics or business.

The University of Chicago Press, 5750 Ellis Ave., Chicago 37, Ill.

AMERICAN ECONOMIC REVIEW

Volume L **December 1960** **Number 5**

ARTICLES

Simulation: A Symposium
- Simulation of Economic Systems ..... G. H. ORCUTT
- Simulation of Industry and Firm ..... MARTIN SHUBIK
- Simulation of Group Behavior..... G. P. E. CLARKSON and H. A. SIMON

Duopoly: Collusion or Warfare? ..... R. L. BISHOP
Income in Relation to Education ..... H. P. MILLER
Progressiveness of a Sales Tax ..... D. G. DAVIES

REVIEW ARTICLE

Employment, Growth, and Price Levels ..... ROMNEY ROBINSON

COMMUNICATIONS

Contributors to the A.E.R. During the 'Fifties
..... F. R. CLEARY and D. J. EDWARDS
The First Two Decades of the American Economic Association
- Comment ..... T. N. CARVER

Competition and Growth—The Lesson of West Germany:
- Comment ..... K. W. ROSKAMP
- Comment..... A. I. MACBEAN and W. G. SHEPHERD
- Comment ..... R. G. OPIE
- Reply ..... EGON SOHMEN

International Trade and Economic Expansion:
- Comment ..... S. T. BEZA

THE AMERICAN ECONOMIC REVIEW, a quarterly, is the official publication of the American Economic Association and is sent to all members. The annual dues are six dollars. Address editorial communications to Dr. Bernard F. Haley, Editor, AMERICAN ECONOMIC REVIEW, Stanford University, Stanford, California. For information concerning other publications and activities of the Association, communicate with the Secretary-Treasurer, Dr. James Washington Bell, American Economic Association, Northwestern University, Evanston, Illinois. Send for information booklet.

PRODUCTIVITY MEASUREMENT REVIEW

Published by the European Productivity Agency

A quarterly review of practical applications and theoretical studies of productivity measurement and interfirm comparison. Recent numbers have contained the following articles:

Investment and Economic Growth: Some Comments ..*R. M. Solow*
Towards a Common Measure of Productivity ..... *H. W. Martin*
A Percentage Measure of Standardisation for use in Interfirm Comparison ..... *S. A. N. Smith-Gavine*
Marketing Productivity and Profitability ..... *Joel Dean*
Business Efficiency: An Interfirm Comparison ..... *L. R. Amey*
Measuring Output in the Public Administration Field ..... *H. D. Lytton*
Interfirm Comparison in Wholesale Trade .... *H. J. A. Van der Post*
Productivity in Retailing ..... *Ben B. Seligman*
Inter-plant Comparisons in Rubber Manufacturing

Each number contains reviews of recent books and articles.

A specimen copy of the *Review* will be forwarded on application to O.E.E.C., 33 rue de Franqueville, Paris 16e, France.

Annual subscription: 1000 Fr. frs.—15s.—U.S. $2.50
or each number: 280 Fr. frs.—4s.—U.S. $0.75

Published in English and French editions

SANKHYA

The Indian Journal of Statistics

Vol. 22 **Parts 3 & 4** **1960**

CONTENTS

The Addition of Random Vectors .....J. B. S. Haldane
A Note on the Estimation of Variance .....Leo A. Goodman
On the Asymptotic Efficiency of Tests and Estimates.....R. R. Bahadur
On Bounded Infinitely Divisible Random Variables
.....Glen Baxter and J. M. Shapiro
A Limit Theorem for Densities .....R. Ranga Rao and V. S. Varadarajan
Tests for Independence and Symmetry in Multivariate Normal Populations
.....J. Roy
On the F-test in the Intrablock Analysis of a Balanced Incomplete Block Design .....Sujit Kumar Mitra
An Admissible Estimate for any Sampling Design .....V. P. Godambe
Non-Null Distribution of the Likelihood-Ratio in Analysis of Dispersion
.....J. Roy
Nonparametric Linear Estimation of Common Median of Symmetrical Populations from Symmetrically Censored Observations.....John E. Walsh
On Some Classification Problems—I .....S. John
On Some Classification Statistics .....S. John
Multivariate Analysis: An Indispensable Statistical Aid in Applied Research .....C. Radhakrishna Rao
Maximum Likelihood Estimation for Positively Regular Markov Chains
..... B. R. Bhat
A Note on the Structure of a Certain Stochastic Model .....S. C. Das
A Short Table of the Hypergeometric Function ${}_0F_1(c; z)$ ..J. Edward Jackson
Approximate Confidence Interval for Linear Functions of Means of K Populations when the Population Variances are not Equal
..... Saibal Kumar Banerjee
Measurement of Attrition in Technical Education .....K. C. Seal
A Note on the Efficiency of Double Sampling for Stratification
..... M. V. Jambunathan
On the Standard Error of the Lorenz Concentration Ratio
..... N. Sreenivasa Iyengar
On a Problem of Estimating Increase in Consumer Demand
..... N. Sreenivasa Iyengar

Subscription Price for Series A and Series B per Volume: U.S.A. and Canada: $10.00; Other Countries: £3; Subscription Price per Volume for both the Series taken together: U.S.A. and Canada: $18.00; Other Country: £5-6s. Subscription is payable to Secretary, Statistical Publishing Society, 204/1 Barrackpore Trunk Road, Calcutta-35, India.

THE INTERNATIONAL ECONOMIC REVIEW is interested primarily in publishing articles on quantitative economics and welcomes contributions of empirical works, as well as those in mathematical economics and statistical theory related to quantitative aspects of economics. Quantitative economics will be broadly interpreted as covering both obviously quantitative studies and ideas that can eventually lead to quantitative studies. In spite of the emphasis on quantitative economics, the editors welcome also contributions to nonquantitative economics.

The Review is published fully in English, three times a year. The annual subscription price is $4.00 per year outside Japan and ¥1,200 per year in Japan; the price of a single copy is $1.50 outside Japan, ¥500 in Japan. Subscriptions outside of Japan should be forwarded to THE INTERNATIONAL ECONOMIC REVIEW, Wharton School of Finance and Commerce, University of Pennsylvania, Philadelphia 4, Pa., U.S.A. Subscriptions in Japan should be forwarded to THE INTERNATIONAL ECONOMIC REVIEW, Kansai Economic Federation, Osaka Building, Soze-cho, Kita-ku, Osaka, Japan.

Manuscripts for publication should be sent to the chief editor or co-editor of the Review.

Editorial communications, books and periodicals for review, and communications respecting advertisements are to be sent to the business office of the Review.

SPRINGER-VERLAG IN WIEN

Linear Programming in Industry

Theory and Applications

An Introduction

By

Sven Danø

Assistant Professor of Economics, University of Copenhagen

In English. With 6 figures. VIII, 120 pages. Gr.-8°. 1960. $ 4.75

The last decade has witnessed a striking development in the field of Operations Research, a branch of science concerned with the application of mathematical and statistical methods to the solution of economic optimization problems in business and other spheres. One of the most important methods of Operations Research is Linear Programming, a mathematical method which has been successfully applied to many practical problems of production, inventory, and transportation planning in industry and agriculture as well as to problems of economic theory.

The book is intended as an introductory text for students of economics as well as for engineers, business managers and others responsible for the organization and planning of industrial operations. It is based on lectures given at several European and American universities and courses held for members of the Association of Danish Engineers.

The book gives a very simple account of the mathematical theory of Linear Programming (with proofs of the basic theorems in an Appendix), describes in detail a large number of industrial applications, and explains the practical methods of numerical solution. Instructive excercises are given throughout the text. The book requires little mathematical proficiency from the reader without sacrificing rigour, and gets further by purely elementary mathematical tools than other existing texts in Linear Programming. Those who have read the book will be capable of detecting, formulating, and solving such linear planning problems as might occur within their respective fields.

OXFORD ECONOMIC PAPERS

Editorial Board: N. H. Leyland (General Editor), J. R. Hicks, C. M. Kennedy, Sir Donald MacDougall, G. B. Richardson, P. P. Streeten (Secretary).

Volume 12 **June 1960** **Number 2**

CONTENTS

Thoughts on the Theory of Capital—The Corfu Conference ..... By J. R. HICKS
Returns to Scale ..... By M. H. PESTON
Israel: Ten Years of Economic Dependence ..... By S. RIEMER
The Working of the Radcliffe Monetary System ..... By A. E. JASAY
Equity and a Capital Gains Tax ..... By R. TURVEY
A Static Interpretation of some Recent Theories of Growth and Distribution ..... By CHARLES KENNEDY
Industrialization in Overpopulated Areas, A Geometric Interpretation of Certain Aspects ..... By F. SETON
Cyclical Fluctuations and Economic Growth .. By LUIGI L. PASINETTI

Oxford Economic Papers is published three times a year in February, June and October. The pre-paid subscription price is 35s. 0d. post free, and the price of a single copy 16s. 0d. net. Orders may be placed with any bookseller or sent direct to the Publisher.

One free specimen copy will be sent on request.

OXFORD UNIVERSITY PRESS

AMEN HOUSE, WARWICK SQUARE, LONDON, E.C.4.

The Quarterly Journal of Economics

Founded 1866

Vol. LXXIV **November 1960** **No. 4**

Utility, Strategy, and Social Decision Rules ..... William Vickrey
A Conservative's Ideas on Economic Reform ..... Thomas Nixon Carver
Market Structure and Uncertainty in Oil and Gas Exploration ..... James W. McKie
International Liquidity and the Balance of Payments of a Reserve-Currency Country ..... Peter B. Kenen
The Erosion of Marshall's Theory of Value ..... Peter Newman
Comment ..... D. H. Robertson
A Theory of Inter-firm Organization ..... Almarin Phillips
Monetary-Fiscal Policy and Economic Growth in an Open Economy ..... David C. Smith

Notes and Discussions

An Alternative Proof of the Factor Price Equalization Theorem .. J. Vanek
Alternative Profit Criteria ..... D. M. Lamberton
State and Regional Payments Mechanisms:
Comment ..... Richard L. Pfister
Reply ..... James C. Ingram
Structural Aspects of Monetary Velocity
Comment ..... William L. Miller
Reply ..... George Garvy
The Politics of Political Economists:
Comment ..... Murray N. Rothbard
Comment ..... A. W. Coats
Reply ..... George J. Stigler

Recent Publications

Edited by Arthur Smithies for the Department of Economics, Harvard University
$1.75 per copy; $6.00 a year; 3 years $17.00; Special Student Rate, $3.00 a year

Order from **THE HARVARD UNIVERSITY PRESS**

Cambridge 38, Massachusetts